

**T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**PANDEMİ DÖNEMİNDE TÜRKİYE’DE, KARGO
FİRMALARINA YAPILAN ŞİKAYETLERİN MAKİNE
ÖĞRENMESİ YÖNTEMLERİ İLE ANALİZİ**

YÜKSEK LİSANS TEZİ

Tolga KUYUCUK

**Enstitü Anabilim Dalı : BİLİŞİM SİSTEMLERİ
MÜHENDİSLİĞİ**
Tez Danışmanı : Dr. Öğr. Üyesi Levent ÇALLI

Haziran 2022

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**PANDEMİ DÖNEMİNDE TÜRKİYE'DE, KARGO
FİRMALARINA YAPILAN ŞİKAYETLERİN MAKİNE
ÖĞRENMESİ YÖNTEMLERİ İLE ANALİZİ**

YÜKSEK LİSANS TEZİ

Tolga KUYUCUK

**Enstitü Anabilim Dalı : BİLİŞİM SİSTEMLERİ
MÜHENDİSLİĞİ**

**Bu tez 17/06/2022 tarihinde aşağıdaki jüri tarafından oybirliği / oyçokluğu ile
kabul edilmiştir.**

Jüri Başkanı

Üye

Üye

BEYAN

Tez içindeki tüm verilerin akademik kurallar çerçevesinde tarafımdan elde edildiğini, görsel ve yazılı tüm bilgi ve sonuçların akademik ve etik kurallara uygun şekilde sunulduğunu, kullanılan verilerde herhangi bir tahrifat yapılmadığını, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezde yer alan verilerin bu üniversite veya başka bir üniversitede herhangi bir tez çalışmasında kullanılmadığını beyan ederim.

Tolga KUYUCUK

TEŞEKKÜR

Yüksek lisans eğitimim boyunca değerli bilgi ve deneyimlerinden yararlandığım, her konuda bilgi ve desteğini almaktan çekinmediğim, araştırmanın planlanmasından yazılmasına kadar tüm aşamalarında yardımlarını esirgemeyen, teşvik eden, aynı titizlikte beni yönlendiren değerli danışman hocam Dr. Öğr. Üyesi Levent ÇALLI'ya teşekkürlerimi sunarım.

Tezimde kullandığım makine öğrenmesi yöntemleri üzerinde bana yardımcı olan yeni fikirler katan ve tez konum ile ilgili kendimi geliştirebilmem için Rusya'da staj olanağı sağlayan değerli Ilya SIGANOV'a teşekkür ederim.

Ayrıca eğitim hayatım boyunca her zaman destek olan, maddi ve manevi yardımlarını esirgemeyen anneme, babama, kardeşime ve bugünlere gelmemde üzerimde emeği bulunan tüm öğretmenlerime teşekkürü bir borç bilirim.

İÇİNDEKİLER

TEŞEKKÜR	i
İÇİNDEKİLER	ii
SİMGELER VE KISALTMALAR LİSTESİ	iv
ŞEKİLLER LİSTESİ	v
TABLolar LİSTESİ	vii
ÖZET	x
SUMMARY	xi
BÖLÜM 1.	
GİRİŞ	1
BÖLÜM 2.	
KAYNAK ARAŞTIRMASI	3
BÖLÜM 3.	
MATERYAL VE YÖNTEM	12
3.1. Materyal	12
3.1.1. Veritabanı	12
3.1.2. Web kazıma modülü.....	12
3.2. Yöntem	13
3.2.1. Doğal dil işleme	14
3.2.2. Etiket ataması.....	14
3.2.2.1. Kelime bazlı etiket ataması	14
3.2.2.2. Anket yöntemi ile etiket ataması	16
3.2.3. Lojistik regresyon	18

3.2.4. One vs rest classifier	18
3.2.5. Vektörleştirme	19
3.2.5.1. Fasttext	19
3.2.5.2. Tfidf vectorizer	19
3.2.5.3. Hashing vectorizer	20
3.2.5.4. Count vectorizer	20
3.3. Analiz	20
3.3.1. Karışıklık matrisi	21
 BÖLÜM 4.	
ARAŞTIRMA BULGULARI	23
4.1. Kelime Bazlı Etiket Ataması ile Eğitilen Model.....	26
4.1.1. Kelime bazlı fasttext yaklaşımı	27
4.1.2. Kelime bazlı tfidf vectorizer yaklaşımı	30
4.1.3. Kelime bazlı count vectorizer yaklaşımı	32
4.1.4. Kelime bazlı hashing vectorizer yaklaşımı	35
4.2. Anket Yöntemiyle Etiket Ataması ile Eğitilen Model	37
4.2.1. Anket yöntemiyle tfidf vectorizer yaklaşımı.....	38
4.2.2. Anket yöntemiyle count vectorizer yaklaşımı	41
4.2.3. Anket yöntemiyle hashing vectorizer yaklaşımı	43
4.3. Kelime Bazlı Etiket Ataması ile Eğitilen Model ve Anket Yöntemi ile Eğitilen Model Sonuçlarının Değerlendirilmesi.....	46
 BÖLÜM 5.	
TARTIŞMA VE SONUÇ	49
 KAYNAKLAR	55
EKLER	59
ÖZGEÇMİŞ	79

SİMGELER VE KISALTMALAR LİSTESİ

CSS	: Cascading Style Sheets
DK	: Dakika
DSÖ	: Dünya Sağlık Örgütü
E-TİCARET	: Elektronik Ticaret
ID	: Identification Number
KNN	: K-Nearest Neighbors
ML	: Machine Learning
NLP	: Natural Language Processing
SVM	: Support Vector Machines
T.C.	: Türkiye Cumhuriyeti
TF	: Term Frequency
TFIDF	: Term Frequency-Inverse Document Frequency
XML	: Extensible Markup Language
XPATH	: XML Path Language

ŞEKİLLER LİSTESİ

Şekil 3.1. Veri analizi süreci	13
Şekil 3.2. Doğal dil işleme adımları	14
Şekil 3.3. Kelime bazlı etiket ataması akış diyagramı.....	16
Şekil 3.4. Şikâyetleri içeren Google Sheets sayfası	16
Şekil 3.5. Şikâyetleri içeren dinamik web sayfası	17
Şekil 3.6. Google Form anket sayfası.....	18
Şekil 4.1. Mart 2020 ve Eylül 2020 tarihleri arasındaki şikâyetlerin zamana göre yoğunluk haritası	24
Şekil 4.2. Türkiye’deki Mart 2020 ve Eylül 2020 tarihleri arasındaki günlük vaka sayıları	24
Şekil 4.3. Türkiye’deki Mart 2020 ve Nisan 2022 tarihleri arasındaki günlük vaka sayıları.....	25
Şekil 4.4. Türkiye’deki Mart 2020 ve Nisan 2022 tarihleri arasındaki toplam vaka sayısı	25
Şekil 4.5. Türkiye’de Mart 2020 ve Nisan 2022 tarihleri arasındaki günlük şikâyet sayısı	26
Şekil 4.6. Şikâyetleri kategorize eden kod bloğu örneği	26
Şekil 4.7. Tfidfvectorizer ile tahminlenen verilerin, kelime bazlı etiket atamalarıyla arasındaki benzerlik oranı	31
Şekil 4.8. Countvectorizer ile tahminlenen verilerin, kelime bazlı etiket atamalarıyla arasındaki benzerlik oranı	34
Şekil 4.9. Hashingvectorizer ile tahminlenen verilerin, kelime bazlı etiket atamalarıyla arasındaki benzerlik oranı	36
Şekil 4.10. Anket yöntemiyle katılımcılardan alınan şikâyet yanıtları	38
Şekil 4.11. Tfidfvectorizer ile tahminlenen verilerin, anket bazlı etiket atamalarıyla arasındaki benzerlik oranı	40

Şekil 4.12. Countvectorizer ile tahminlenen verilerin, anket bazlı etiket atamalarıyla arasındaki benzerlik oranı	42
Şekil 4.13. Hashingvectorizer ile tahminlenen verilerin, anket bazlı etiket atamalarıyla arasındaki benzerlik oranı	45



TABLolar LİSTESİ

Tablo 2.1. Müşteri şikâyetleri yönetimi konusundaki literatür özeti	9
Tablo 3.1. Kargo firmalarına yapılan bir şikâyet içerik örneği	15
Tablo 3.2. Count vectorizer örneği	20
Tablo 3.3. Karışıklık matrisi	21
Tablo 4.1. Tüm içeriklerde en çok geçen ilk 10 kelime	23
Tablo 4.2. Şikâyetlerin kelime ve kelime öbeği bazlı kategorize edilmesi	27
Tablo 4.3. Fasttext modelde “corona” kelime vektörleri	27
Tablo 4.4. Fasttext modelde “corona”, “covid”, “araba” ve “eldiven” kelimelerine yakın kelimeler	28
Tablo 4.5. Fasttext orjinal örneği	29
Tablo 4.6. Fasttext multilabel kargo şikâyet örneği	29
Tablo 4.7. 3000 eğitim verisinde ilk 10 örnek üzerinde TfidfVectorizer tahmin oranları	30
Tablo 4.8. 3000 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak TfidfVectorizer tahmin oranları.....	31
Tablo 4.9. Tfidfvectorizer ile kelime bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	32
Tablo 4.10. 3000 eğitim verisinde ilk 10 örnek üzerinde CountVectorizer tahmin oranları.....	33
Tablo 4.11. 3000 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak CountVectorizer tahmin oranları.....	33
Tablo 4.12. CountVectorizer ile kelime bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	34
Tablo 4.13. 3000 eğitim verisinde ilk 10 örnek üzerinde HashingVectorizer tahmin oranları.....	35

Tablo 4.14. 3000 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak HashingVectorizer tahmin oranları.....	36
Tablo 4.15. HashingVectorizer ile kelime bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	37
Tablo 4.16. Kelime bazlı etiket ataması karışıklık matrisi	37
Tablo 4.17. 1887 eğitim verisinde ilk 10 örnek üzerinde TfidfVectorizer tahmin oranları	39
Tablo 4.18. 1887 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak TfidfVectorizer tahmin oranları	39
Tablo 4.19. Tfidfvectorizer ile anket bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	40
Tablo 4.20. 1887 eğitim verisinde ilk 10 örnek üzerinde CountVectorizer tahmin oranları	41
Tablo 4.21. 1887 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak CountVectorizer tahmin oranları	42
Tablo 4.22. CountVectorizer ile anket bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	43
Tablo 4.23. 1887 eğitim verisinde ilk 10 örnek üzerinde HashingVectorizer tahmin oranları	44
Tablo 4.24. 1887 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak HashingVectorizer tahmin oranları	44
Tablo 4.25. HashingVectorizer ile anket bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	45
Tablo 4.26. Toplam şikâyet verilerinin eğitilen yöntemlerine göre kategorilere dağılımı	46
Tablo 4.27. Toplam şikâyet sayısının kargo firmaları bazında dağılımı	47
Tablo 4.28. Toplam şikâyet verisi değerlerinin kargo firmaları bazında kategorilere göre dağılımı	47
Tablo 4.29. Toplam şikâyet verisi değerlerinin kargo firmaları bazında kategorilere göre yüzdelik dağılımı	48
Tablo 4.30. Anket yöntemiyle etiket ataması karışıklık matrisi	48

Tablo 5.1. Etiketleme ve vektörlendirme yöntemlerine göre her bir kategori için tahminlenen şikâyet sayıları	50
Tablo 5.2. Tahminlenen şikâyetlerin doğruluk oranları	50
Tablo 5.3. Tahminlere göre hazırlanan karışıklık matrisleri	51



ÖZET

Anahtar Kelimeler: Kargo şirketleri, covid-19, multilabel sınıflandırma, web kazıma, müşteri şikâyetleri, metin madenciliği

Bu çalışma, Türkiye'de hizmet sektörü içerisinde önemli bir yer tutan lojistik, nakliye ve kargo işletmelerinin Covid-19 salgını sırasında süreçlerini nasıl yönettiklerini keşfetmeye çalışmakta ve olumsuz sonuçlara yönelik çözümler önermektedir.

Bu çalışma içerisinde kullanılan veri, pandeminin başladığı tarihten çalışmanın yürütüldüğü tarihe kadar (11.03.2020 – 30.09.2021) online bir şikâyet yönetimi web sitesinden taşımacılık ve kargo sektöründeki işletmeler ile ilgili yapılmış olan ve içerisinde Covid-19 salgını ile ilgili kelimeleri bulunduran şikâyetler Python programlama dilinde web kazıma için kullanılan Scrapy modülü ile toplanmıştır.

Toplanan şikâyet verilerinin ilk olarak hangi konular üzerinden en çok şikâyet aldığını belirleyerek: “Gecikti veya Dağıtıma Çıkmadı”, “Evde yok notu düşüldü veya Kapıya Getirilmedi”, “Telefonlara Cevap Verilmedi”, “İade Süreci”, “Teslim Alınmadı veya Teslim Edilmedi” ve “Hijyen Kurallarına Uyulmadı” olmak üzere birbirinden farklı 6 kategoriye ait olabileceği saptanmıştır. Ardından, elde edilen tüm içeriğin bir kısmı kategorilendirilerek, belirlenmiş olan 6 kategoriye göre sınıflandırılmış ve eğitim verisi olarak kullanılmıştır. Eğitim verisinin etiketlenmesinde kelime bazlı ve anket yöntemiyle olmak üzere iki farklı etiket ataması yöntemi kullanılmıştır.

Kalan içerik, eğitim verisinden elde edilen analizlere göre multilabel sınıflandırma algoritmalarından faydalanılarak sınıflandırılmıştır.

Araştırmada elde edilen bulgulara göre en çok şikâyetin kargoların dağıtımı sırasında yaşandığı ve gecikme yaşanan vaka sayısının yüksek olduğu sonucuna varılmıştır.

DURING PANDEMIC PERIOD IN TURKEY, ANALYSIS THE COMPLAINTS TO THE CARGO COMPANIES BY MACHINE LEARNING METHODS

SUMMARY

Keywords: Covid-19, web scraping, cargo companies, customer complaints, multilabel classification, text mining

This study attempts to discover how cargo, transportation, and logistics firms, which have a considerable market share in Turkey's service sector, manage their processes during the Covid-19 outbreak and suggests solutions to the adverse outcomes.

The data used in the research include complaints made for firms in the cargo sector from an online complaint management website (sikayetvar.com) from the beginning of the pandemic to the date of the research (11.03.2020 – 30.09.2021), which contains the words related to the pandemic were collected with Python language and Scrapy module web scraping methods.

The complaints obtained were evaluated with qualitative methods, and the six most observed topics were identified. These topics are as follows; delayed or not delivered, not at home or not delivered”, phones not answered, return process, not received or delivered, and hygiene rules not followed.

Complaints were classified by using multilabel classification algorithms according to the analyzes obtained from the training data created according to the determined topics. According to the findings, it was observed that the most complaints were experienced during the delivery process of parcels, and the number of delayed issues was the highest.

BÖLÜM 1. GİRİŞ

Çalışmanın Amacı: Bu çalışma, Türkiye'de hizmet sektörü içerisinde önemli bir yer tutan lojistik, nakliye ve kargo işletmelerinin Covid-19 salgını sırasında süreçlerini nasıl yönettiklerini keşfetmeye çalışmakta ve olumsuz sonuçlara yönelik çözümler önermektedir.

Çalışmanın Önemi: Bu tez çalışması kapsamında pandemi dönemi içerisinde kargo firmalarına yapılan şikâyetlerin nedenleri tespit edilmiştir. Çalışmada kullanılan eğitim verisinin etiketlenmesinde kelime bazlı ve anket yöntemiyle olmak üzere iki farklı etiket ataması yöntemi kullanılmıştır. Böylece literatüre veri etiketleme konusunda katkı sağlanmak amaçlanmıştır.

İlk olarak 2019 yılının son günlerinde Çin'in Wuhan kentinde ortaya çıkan Covid-19 insandan insana geçebilmekte ve geçtiği insanı birkaç gün içerisinde enfekte edebilme özelliğiyle dünyada hızlı bir biçimde yayılmıştır. Türkiye'de ilk vaka 11 Mart 2020 de görülmüş ve o tarihten 2022 yılı ilk çeyreğine kadar (Resmigazete.gov.tr, 2022) virüsün yayılmasını önlemek için sürekli olarak önlemler alınmıştır. Virüsün tükürük ve damlacık yoluyla bulaştığı tespit edilmiştir. Bu yüzden Türkiye'de alınan önlemlerden biri de dışarıda maske takma zorunluluğu olmuştur.

Her vatandaş gibi hizmet sektörü ve kurumsal firma çalışanlarının da bu önlemlere uymaları hem kendi sağlıkları hem de iletişime geçtikleri müşterilerinin sağlıkları açısından hayati derecede önemlidir.

Hizmet sektöründeki firmalar, müşterilerinin memnuniyetlerini ve şikâyetlerini, ürün veya hizmet aldıkları firmalara iletmesiyle, bu değerlendirmeleri doğru algıladıkları takdirde kendilerini geliştirmekte ve günden güne daha iyi hizmet verir hale

gelmektedirler. Günümüzde firmalar ile müşterileri arasındaki iletişim daha önce hiç olmadığı kadar yüksektir ve bu iletişim her geçen günde de artmaktadır. Bunun sebebi iletişim kanallarının teknoloji ile paralel olarak doğru oranda artması ve iletişimin kolaylaşmasıdır. Mevcut iletişim kanallarından en etkin olanlarından biri ise internet siteleridir. Müşteriler bazı internet sitelerine ürün veya hizmet aldıkları firmalar hakkındaki görüşlerini yazarak bu firmalar ile iletişim kurabilmekte ve firmalardan dönüş alabilmektedirler.

Türkiye’de yaygın olarak kullanılan ve müşteriler ile firmalar arasında köprü görevi gören online bir şikâyet yönetimi platformundan elde edilen verilere göre, Covid-19 salgının başladığı günden bu çalışmanın yapıldığı tarihe kadar en çok şikâyetin kargo ve taşımacılık firmalarına yapıldığı tespit edilmiştir. Kargo firmalarına gelen bu şikâyetler çok çeşitli olmakla birlikte özellikle bazı konularda daha yoğun şikâyetlerin yapıldığı görülmüştür.

Bu çalışmada, online bir şikâyet yönetimi platformu üzerinden Türkiye’deki kargo ve taşımacılık firmalarına yapılan şikâyetler değerlendirilerek pandemi dönemi içerisinde lojistik sektörü ile ilgili en önemli sorunlar multilabel algoritma yöntemleri ile tespit edilerek sınıflandırılmış ve makine öğrenmesi yöntemlerinden faydalanılarak sonuçlar üretilmiştir. Aynı zamanda ileride benzer sorunların yaşanma olasılığının azaltılması amacıyla, elde edilen sonuçlar üzerinden bu sorunların çözümüne yönelik ilgili firmalara durum tespitleri ve öneriler sunulmuştur.

Çalışma içerisinde kullanılan veriler Türkiye’de faaliyet gösteren 5 büyük kargo firmasını kapsamaktadır. Kargo firmalarının isimleri gizlenerek: A, B, C, D ve E Firmaları şeklinde ifade edilmiştir.

BÖLÜM 2. KAYNAK ARAŞTIRMASI

Bu bölümde; Covid-19 ve pandemi dönemi, e-ticaret kavramı, Türkiye'de ve dünyada e-ticaretin durumu, kargo ve Lojistik firmalarının e-ticaret içindeki yeri ve Türkiye'deki kargo firmalarına yönelik müşteri şikâyetleri konularında literatürde bulunan kaynaklardan derlenen bilgiler aktarılmıştır.

Aralık 2019'un son günlerinde Çin'in Wuhan kentinde nedeni bilinmeyen bir grup zatürre vakası ortaya çıkmış ve bu konuda bir araştırma başlatılmıştır. Araştırma sonuçlarına göre, bu enfeksiyonun kaynağının Wuhan Güney Çin Deniz Ürünleri Pazarı ile ilgili olabileceği belirtilmiştir. Ocak ayında Çin hükümeti bu durumu DSÖ'ye bildirmiş ve bu salgına neden olan patojen 7 Ocak 2020'de yeni bir tip korona virüs olarak tanımlanmış ve Covid-19 olarak adlandırılmıştır (Yangping, 2020). Vaka sayısının 23 Ocak 2020'ye kadar hızlı bir şekilde katlanarak büyümesinin ardından salgın diğer ülkelere de sıçramış ve dünya çapında büyük bir yankı uyandırmıştır (ZU ve ark., 2020).

Türkiye'de de ilk Covid-19 vakası 11 Mart tarihinde görülmüştür. DSÖ aynı tarihte, yeni ortaya çıkan bu virüsün bir pandemi olduğunu bildirmiştir (Budak ve Korkmaz, 2020). Covid-19 salgını sebebiyle, dünya çapında milyonlarca insan enfekte olmuştur. 11 Mayıs 2022 tarihi itibarıyla toplamda 518,501,698 vaka görülmüştür. Bunların 38,782,962'si hâlen aktif vaka iken 479,718,736'sı sonuçlandırılmış vakalardır. Sonuçlandırılmış vakalar arasından 6,280,299 kişi hayatını kaybetmiş ve 473,438,437 kişi ise iyileşmiş olarak rapor edilmiştir. Enfekte olan vakaların %1'i hayatını kaybetmiştir (Worldometers.info, 2022). Bu hastalıktan en sık etkilenenler ise orta yaş ve üzerinde olan bireyler olmaktadır. Covid-19 olduğu kesinleşmiş hastalar üzerinde yapılan çalışmalara göre hastalığın yoğunluğunun ortanca yaşlar (49-56) arasında değiştiği görülmüştür. Çin verilerine dayalı olarak yapılan bir çalışmada, Covid-19 tanısı konularak hastaneye yatma oranlarının yaşa bağlı olarak

arttığı, genç yaş (20-29) için %1, orta yaş üzeri (50-59) için %4 ve 80 yaşından büyük bireyler için %18 oranında olduğu raporlanmıştır (Altın, 2020).

Covid-19 sebebiyle Türkiye’de; iş yeri açma-kapama saatleri, toplu taşımada kişi sayısı sınırı, kişiler arası mesafe sınırı (sosyal mesafe), maske takma zorunluluğu, HES kodu zorunluluğu, sokağa çıkma kısıtlaması gibi konularda çeşitli önlemler alınmıştır. Sağlık için yapılan bu hamle ekonomi açısından olumsuz etkiler oluşturarak hizmet sektöründeki kurumsal firmaların çalışmalarını güç hâle getirmiştir. Bu durumun başlıca nedeni ise tüketicilerin pandemi önlemleri sebebiyle dışarıya çıkmalarının kısıtlanmış olmasına bağlı, ihtiyaç duyduğu hizmet ve ürünlere ulaşamamasıdır. Bu durum sebebiyle tüketiciler geleneksel yollar yerine internet yoluyla satın alma yoluna gitmiştir. Müşterilerine ulaşmak için firmaların büyük bir kısmı pandemi döneminde e-ticarete yönelmiş ve zaten e-ticaret faaliyetleri yürüten firmalarda bu alandaki hacimlerini artırmıştır. Elektronik ticaret ya da kısaca e-ticaret, ticaretin elektronik ortamda yapılması olarak tanımlanmaktadır.

E-ticarete olan yönelim, şirketlerin müşterilerine daha fazla seçenek sunmasına olanak tanımıştır. E-ticaret sitelerinde müşteriler aradıkları ürünleri belirli kategoriler arasından seçme şansına sahiptir. Bu da müşterilerin seçimlerini daraltmalarına ve satın alma kararlarında kendilerini daha güvenli hissetmelerine sebep olmaktadır (Schafer ve ark., 1999).

E-ticaret, Türkiye’de ilk olarak 1997 yılında ülkemizde elektronik bir ticaret ağı oluşturmak üzere Bilim ve Teknoloji Yüksek Kurulu’nun bu konuda toplantılar düzenlemesiyle gündeme gelmiştir. Yıllar içerisinde çok sayıda çalışma ve kanun düzenlemeleri sonucunda günümüzde Türkiye’de e-ticaret ağı genişlemiş ve yaygınlaşmıştır (Demirdöğmez ve ark., 2018). Pandemi, dünya çapında tüm ülkelerde e-ticaret sektörünün büyüme sürecini hızlandırmıştır. 2019 yılında belirli bir ülkede e-ticaret konusunda gelişmişlik seviyesi ne kadar düşükse, Covid-19 salgını sırasında e-ticaretin büyüme oranının o kadar yüksek olduğu görülmüştür. Bu, çok düşük e-ticaret hacmine sahip ülkelerin bu konuda hızla gelişmekte olduğu anlamına gelmektedir (Alfonso ve ark., 2020).

2020 yılında Türkiye'nin e-ticaret hacminin 2019 yılına göre yüzde 66 oranında artarak 136 milyar liradan 226,2 milyar liraya ulaşmış, 2021 yılında ise 2020 yılına göre yüzde 69 oranında bir artışla 381,5 milyar lira olarak gerçekleşmiştir (Eticaret.gov.tr, 2022). Pandemi nedeniyle çevrimiçi alışverişe yönelimdeki artış, çevrimiçi pazarların hem gelirlerini hem de maliyetlerini artırmıştır. Global çapta pazar hakimiyeti yüksek e-ticaret firmaları için 2020'nin ilk yarısında elde ettikleri gelirler yükselmiştir. Amazon firması, aniden oluşan fazla iş hacmi nedeniyle salgın sırasında 175.000 yeni çalışanı işe almış ve test laboratuvarları ile termal kameralar da dahil olmak üzere koruyucu ekipmanlar için tahmini olarak 4 milyar dolar harcamıştır (Alfonso ve ark., 2020). Tran (2021), pandemi döneminde gerçekleştirdiği çalışmasında firmalara stratejik olarak e-ticaret platformları oluşturmalarını ve tedarikler için çevrimdışı yöntemlerle birlikte çalışmalarını önermiştir. Buna neden olarak, tüketicilerin bulaşıcı hastalıklardan kaçınmak için çevrimiçi kaynaklara yöneldiğini ve giderek daha fazla sürdürülebilir tüketim davranışları sergiliyor olmalarını göstermiştir.

Aktaş ve ark., (2020), Trendyol.com e-ticaret sitesinde paylaşılmış Türkçe yorumlar üzerine yapmış oldukları araştırmalarında, pandemi sürecinde müşterilerin online olarak satın aldıkları ürünlere yapmış oldukları değerlendirmelerinde ürün memnuniyetinden daha çok, ürün kullanılabilirliği ve ürün özelliklerinden bahsetmiş olduklarını belirtmişlerdir. E-ticaret sitelerinde var olan sipariş yoğunluğunun giderek artması, bu alandaki sorunları da büyük ölçüde artırmaktadır. Kargo çalışanlarının sağlığı ya da tüketicilerin satın aldıkları ürünlerin zamanında ve sorunsuz bir şekilde ulaşması gibi konular bu süreçte oldukça önemlidir. Ürünlerin kargoyla teslimatında ise yoğunluğa bağlı olarak büyük aksamlar yaşanmaktadır. Zamanında teslim edilemeyen çok sayıda ürün olduğu da bilinmektedir (Demiröğmez ve ark., 2020). Türkiye'de kurumsal firmalarda yönetim faaliyetlerini yürüten yöneticiler ve KOBİ'ler üzerinde uygulanan bir anket çalışmasına göre, Covid-19 pandemisi döneminde KOBİ'lerin %38'i kargo sıkıntıları yaşadıklarını belirtmiştir. Aynı çalışmada sektörde beklenen en büyük risk olarak e-ticaret sektöründeki büyümeyi kargo operasyonlarının karşılayamaması, e-ticaret

büyümesini limitlendirmesi ve buna bağlı olarak da müşteri memnuniyetsizliğini artırması olarak gösterilmektedir (Ey.com, 2020).

Günümüzde var olan ürün çeşitliliği ve gelişen teknolojinin getirdiği kolaylıklar sonucu oluşan beklentiler, hem müşteri ve işletme arasındaki ilişkileri hem de işletmelerin birbirleri ile aralarında olan ilişkilerini farklı boyutlara taşımıştır. İşletmeler müşterilerinin ihtiyaçlarını daha hızlı bir şekilde karşılamak için stoklarla çalışılan standart yaklaşımlarını önemli ölçüde değiştirmiştir. Değişimlere bağlı olarak hızlı ve güvenli gönderilere imkân sağlayabilen kargo ve taşımacılık hizmetleri hem bireyler hem de kurumlar tarafından daha fazla kullanılır hâle gelmiştir (Songur ve Büyükkeklik, 2016).

Sabuhari ve Irawanto (2020), kargo firmalarını daha verimli bir hale getirmek için öncelikle çalışanların performansının artırılmasının gerekliliğini savunmaktadır. Bunun için yapılan öneri ise stratejik insan kaynakları yönetimidir. Nikolaeva ve ark. (2020), çalışmalarında lojistik sektöründe teknoloji ve dijitalleşmenin önemi ve kargo taşımacılığının daha iyi hale getirilmesi konusundaki rolünden bahsetmişlerdir. Jimenez ve ark. (2018), Barranquilla (Kolombiya) şehrinde faaliyet gösteren kargo ve taşımacılık firmaları üzerinde yaptıkları çalışmalarında, karayolu kargo taşımacılığının, lojistik zincirinde ekonomik ve sosyal nedenleriyle ulusal ve küresel düzeyde büyük önem arz ettiğinden bahsetmişlerdir. Ancak bu sektörün üretkenliğini daha iyi seviyelere çıkartmak, diğer bölge ve ülkelerle rekabet edebilmek ve bu hedeflere ulaşabilmek adına operasyonel eksiklikleri giderebilmek için Entegre Yönetim Sistemleri'nin kullanılmasının gerekliliğini savunmuşlardır. Entegre Yönetim Sistemleri'nin kaliteli bir şekilde uygulanabilmesini sağlamak için ise, başlangıç koşullarının tespiti, çevre yönetimi, iş sağlığı ve güvenliği ve lojistik operasyonlarda güvenlik kriterlerine dayalı bir metodoloji önermişlerdir.

Lisińska-Kuśnierz ve Gajewska (2014), lojistik hizmetlerin hizmet kalitesinin ağırlıklı olarak müşteri memnuniyetinin ölçülerek değerlendirildiğini ve müşteri memnuniyetinin ölçülmesinin, kuruluşların nasıl faaliyet gösterdiği ve müşteri ihtiyaçlarının etkin bir şekilde nasıl karşılanacağı hakkında bilgi sağladığını

belirtmişlerdir. Çalışmada, soğutmalı taşımacılık hizmeti alan firmalara anket yoluyla taşımacılık hizmetine bağlı memnuniyet durumu sorgulanmıştır. Sonuçlara göre; teslimatın zamanında yapılması, teslimatın eksiksizliği ve teslimatın doğru olması gibi konular kalite hizmetlerinin değerlendirilmesinde ön plana çıkmıştır.

Türkiye’de kargo sektörü ve müşteri memnuniyeti hakkında yapılan çalışmalara bakıldığında; Deniz ve Gödekmerdan (2011), yapmış oldukları bir araştırmada, Erzurum ilinde faaliyet gösteren kargo ve taşımacılık firmalarından hizmet alan müşteriler arasından 400 kişiye yüz yüze anket uygulayarak, müşterilerin kargo firmalarının sundukları hizmetlerine yönelik tutum ve düşüncelerini belirlemeye çalışmıştır. Cevaplayıcıların, kargo gönderirken en çok karşılaştıkları problemin gecikmeler olduğu belirlenmiştir. Onurlubaş ve Gümüş (2020), yaptıkları çalışmalarında katılımcıların hizmet kalitesi algılarını 5 boyutta toplamışlardır: Yanıt verilebilirlik, Güvenirlik, Empati, Güvence ve Fiziksel görünüm. Tespit edilen bu boyutlar katılımcıların hizmet kalitesi ile ilgili algılarının yaklaşık %67’sini açıklamaktadır. Taşımacılık faaliyetlerinde müşterilerinin memnuniyetini sağlama konusunda sorun yaşayan firmalar, sorunların çokluğu ve büyüklüğüyle orantılı olarak müşterilerinden geri dönüş ve şikâyet bildirimleri almaktadırlar. Müşteri şikâyetlerini iyi yönetebilen firmalar, müşterilerini kaybetmeyi önlemekle kalmayıp müşterilerinin güvenlerini tazelediklerinde daha da sadık müşteriler edinmeyi başarabilmektedirler.

Müşteri şikâyetleri yönetimi konusunda literatüre bakıldığında; Cho ve ark. (2002), çalışmalarında işletmelerin müşteri şikâyeti davranışını etkili bir şekilde nasıl ele aldığını görmek mümkündür. Bu çalışmada, üç farklı e-ticaret müşteri hizmetleri merkezinden gelen 1000 müşteri şikâyeti ve çevrimiçi geri bildirim sistemlerinden gelen 500 adet şikâyet analiz edilmiştir. Araştırma bulguları, mükemmel çevrimiçi müşteri hizmetleri sağlamak, müşterilerin taleplerine / şikâyetlerine hızlı yanıt vermek ve söz konusu ürün kategorisine uygun stratejiler kullanmak olarak elde edilmiştir. Sonuçta şikâyet yönetimini ciddiye almanın ürün/hizmet kalitesine, web sitesi tasarımına ve işletme politikalarına etki ettiği anlaşılmış ve çevrimiçi şikâyet yönetimine yönelik çalışmaların, müşteri sadakatini sağlamayı kolaylaştığı tespit

edilmiştir. Davidow (2003) makalesindeki amacı, özellikle bir müşteri şikâyetine verilen organizasyonel yanıtın şikâyet sonrasında müşteri davranışını nasıl etkilediğine odaklanarak, şikâyet yönetimi alanındaki mevcut araştırmaları özetlemek olarak belirtmiştir. Bu organizasyonel yanıtları altı ayrı boyuta (zamanlama, kolaylaştırma, düzeltme, özür, güvenilirlik ve özen) ayıran ve sırayla her bir boyuta derinlemesine bakılan bir model çerçevesi sunulmuştur. Firmaların müşteri şikâyetlerine geri dönüşlerinde adil bir tutum sergilemelerinin, müşterilerinin davranışları üzerinde büyük bir etkiye sahip olduğu anlaşılmıştır.

Bell ve Luddington (2006) çalışmalarında, müşteri şikâyetleri ile servis personelinin müşteri hizmetlerine bağlılığı arasındaki ilişkiyi araştırmaktadır. 124 mağazalı ulusal bir perakende zincirinde 432 perakende hizmet personeli anketinden elde edilen verileri kullanarak, müşteri şikâyetlerinin, hizmet personelinin müşteri hizmetlerine bağlılığı ile önemli derecede ve olumsuz yönde ilişkili olduğunu bulmuşlardır. Daha yüksek seviyelerdeki hizmet çalışanlarının ise müşteriler ile pozitif ilişkiler kurarak bu olumsuz ilişkiyi önemli ölçüde azalttığı görülmüştür.

Alabay (2012)'a göre şikâyetler sayesinde işletmeler, ürün ve hizmetlerini müşterilerinin beklentilerini giderecek şekilde geliştirip değiştirebileceklerdir. Müşterinin mal ve hizmeti ilk satın almasında tatmin olması ne kadar önemliyse, şikâyeti sonrasındaki tatmin olması da bir o kadar önemlidir. Bu çalışmada müşteri memnuniyetini sağlayan unsurlar incelenmiş ve daha sonra şikâyet faktörleri, amaçları, araçları, süreci ve yönetimi konuları ele alınmıştır. Sonuca göre, terk etmiş müşteriye geri döndürmenin maliyetinin normal yollarla müşteri edinmeye göre 6 kat fazla olduğunun görülmesiyle var olan müşteriye tutmanın ekonomik açıdan önemi anlaşılmıştır. Erdem ve Akolaş (2020), lojistik yönetiminin işletmelere sağlayabileceği faydaları göz önünde bulundurarak işletmelerin bu konuya önem vermeleri gerektiğini belirtmişlerdir (Tablo 2.1.).

Araştırma kapsamında, lojistiğin işletmeler açısından müşteri memnuniyetini ölçme konusunda demografik değişkenler ele alınmıştır. Bulgulara göre müşteri memnuniyeti; cinsiyet, yaş, eğitim durumu, aylık gelir, işletmeyi tercih sebebi ve

işletme ile çalışma sayısı gibi konulara göre farklılık göstermektedir. Sonuç olarak ise işletmelerin mevcut müşterilerini ellerinde tutması ve müşterilerinin sadakatini geliştirmelerinin, yeni müşterilere hitap etmelerinden daha önemli olduğu anlaşılmıştır.

Tablo 2.1. Müşteri şikâyetleri yönetimi konusundaki literatür özeti

Yazar	Değişkenler	Sonuç
Cho ve ark. (2002)	Üç farklı e-ticaret müşteri hizmetleri merkezinden gelen 1000 adet müşteri şikâyeti ve çevrimiçi geri bildirim sistemlerinden gelen 500 adet şikâyet analiz edilmiştir.	Şikâyet yönetimini ciddiye almanın ürün/hizmet kalitesine, web sitesi tasarımına ve işletme politikalarına etki ettiği anlaşılmış ve çevrimiçi şikâyet yönetimine yönelik çalışmaların, müşteri sadakatini sağlamayı kolaylaştığı tespit edilmiştir.
Davidow (2003)	Müşteri şikâyetlerine firmaların verdikleri yanıtları altı ayrı boyuta (zamanlama, kolaylaştırma, düzeltme, özür, güvenilirlik ve özen) ayıran ve sırayla her bir boyuta derinlemesine bakılan bir model çerçevesi sunulmuştur.	Firmaların müşteri şikâyetlerine geri dönüşlerinde adil bir tutum sergilemelerinin, müşterilerinin davranışları üzerinde büyük bir etkiye sahip olduğu anlaşılmıştır.
Bell ve Luddington (2006)	124 mağazalı ulusal bir perakende zincirinde 432 perakende hizmet personeli anketinden elde edilen veriler kullanılarak, müşterilerin şikâyetlerinin hizmet personelinin tutumuna göre olumlu ve olumsuz sonuçları araştırılmıştır.	Hizmet personelinin müşteri hizmetlerine bağlılığı ile önemli derecede ve olumsuz yönde ilişkili olduğu bulunmuş, ancak daha yüksek seviyelerdeki hizmet çalışanlarının müşteriler ile pozitif ilişkiler kurarak bu olumsuz ilişkiyi önemli ölçüde azalttığı görülmüştür.
Alabay (2012)	Müşteri memnuniyetini sağlayan unsurlar incelenmiş ve daha sonra şikâyet faktörleri, amaçları, araçları, süreci ve yönetimi konuları ele alınmıştır.	Terk etmiş müşteriyi geri döndürmenin maliyetinin normal yollarla müşteri edinmeye göre 6 kat fazla olduğunun görülmesiyle var olan müşteriyi tutmanın ekonomik açıdan önemi anlaşılmıştır.
Erdem ve Akolaş (2020)	Lojistiğin işletmeler açısından müşteri memnuniyetini ölçme konusunda demografik değişkenler ele alınmıştır.	İşletmelerin mevcut müşterilerini ellerinde tutması ve müşterilerinin sadakatini geliştirmelerinin, yeni müşterilere hitap etmelerinden daha önemli olduğu anlaşılmıştır.

Pandemiye bağılı olarak kargo ve taşımacılık firmalarıyla sorunlar yaşayan kullanıcıların önemli bir kısmı, yaşadıkları bu sorunları online şikâyet sitelerine yazarak çözüm aramaktadır. Kullanıcıların genel olarak yaşadıkları sorunların hangi konular üzerinde yoğunlaştığını anlamamanın zorluğu, şikâyet sayısının yoğunluğuyla doğru orantılı olarak değişmektedir. Şikâyet hacminin yoğun olduğu bir sektörde her bir şikâyetin incelenmesi, daha fazla insan gücü ve zekasını kullanma durumu gerektireceğinden firmaların bütçeleri için zorlayıcı olabilmektedir. Bunun yerine yalnızca belli sayıda bir şikâyetin incelenmesi ve bu incelemelerin makine öğrenmesi algoritmalarıyla işlenerek daha sonra gelecek olan şikâyetlerin daha çabuk ve düşük maliyetle kategorize edilebilmesine olanak tanıyacaktır.

Makine öğrenmesi kavramı, ağırlıklı olarak dijital dünyada son derece büyük ilgi gören ve dijitalleştirme çözümlerinin de önemli bir bileşeni olan Yapay Zekânın bir alanıdır. Eğitim verisi türlerine ve kategorilerine bağılı olarak, "denetimli öğrenme", "denetimsiz öğrenme", "yarı denetimli öğrenme" ve "pekiştirmeli öğrenme" gibi tekniklerden faydalanılır (Ray, 2019). Makine öğrenmesi aslında çevredeki ortamdan öğrenme yaparak bir bakıma insan zekasını taklit etmek için tasarlanmış olan ve sürekli gelişen bir hesaplama algoritmasıdır (Naqa ve Murphy, 2015). Çoğunlukla her veri problemine uygulanabilir ve yaygın olarak kullanılan makine öğrenmesi algoritmalarına örnek olarak; Lineer Regresyon, Lojistik Regresyon, Karar Ağacı, SVM, Naive Bayes, kNN, K-Means, Random Forest, Boyut İndirgeme Algoritmaları, Gradyan Artırma Algoritmaları verilebilir (Analyticsvidhya.com, 2021).

Metin madenciliği kavramı ise, farklı yazılı kaynaklardan otomatik olarak bilgi çıkartarak yeni ve önceden bilinmeyen bilgilerin bilgisayar tarafından keşfedilmesidir (Kwartler, 2017). Metin madenciliğinde en önemli unsurlardan biri olan doğal dil işleme, bir dizi görev ya da uygulama için insan benzeri dil işlemeyi başarmak amacı çerçevesinde doğal olarak oluşmuş olan metinleri bir veya daha fazla dil analizi düzeyinde analiz etmek ve temsil etmek için teorik olarak motive edilmiş bir hesaplama tekniğidir (Wilks, 1996). Metin madenciliği, bilgi alma, metin analizi, bilgi çıkarma, kümeleme, sınıflandırma, görselleştirme, veritabanı

teknolojisi, makine öğrenimi ve veri madenciliğini de içeren çok disiplinli bir alandır (Tan, 1999).

Sınıflandırma problemlerinin birçoğu, her bir sınıfı bir örnek veya durumla ilişkilendirir. Ancak, her bir örneğin bir veya daha fazla sınıfla ilişkilendirilebildiği birçok sınıflandırma görevi vardır. Bu problem grubu da Multilabel Classification (Çok Etiketli Sınıflandırma) olarak bilinen bir alanı temsil eder. Yani Multilabel Classification, her belgenin birden fazla sınıfa atanabildiği belgelerin sınıflandırılmasıdır (de Carvalho ve Freitas, 2009). Python dili, makine öğrenmesi kütüphaneleriyle birlikte, endüstri için model oluşturma ve araştırmacılar için yeni yöntemler geliştirme konusunda önde gelen teknolojilerden biri haline gelmiştir. Scikit-multilearn, Multilabel Classification'ı gerçekleştirmek için gerekli bir Python kütüphanesidir. Bu kütüphane, popüler Multilabel Classification yöntemlerinin Python diliyle yazılmış uygulamalarda kullanılmasını sağlar (Szyma ve Kajdanowicz, 2016).

BÖLÜM 3. MATERYAL VE YÖNTEM

3.1. Materyal

Çalışmada kullanılan veriler online bir şikâyet yönetimi platformu olan sikayetvar.com web sitesi üzerinden Scrapy modülü kullanılarak Python dilinde yazılan kodlar yardımıyla elde edilmiş ve veritabanına kaydedilmiştir. Şikayetvar sitesi, müşteriler ve markalar arasında köprü görevini üstlenen bir web platformudur (Sikayetvar.com, 2022). Araştırmanın çalışma grubunu, Türkiye’de hizmet veren kurumsal kargo şirketlerinden hizmet alıp, aldıkları hizmetten memnun kalmayarak ve bu memnuniyetsizliklerini pandemi ile alakalı sebeplere dayandırarak sikayetvar.com web sitesi üzerinden dile getiren 16333 kişi oluşturmaktadır.

3.1.1. Veritabanı

Şubat 2020 ile Eylül 2021 tarihleri arasında sikayetvar.com web sitesi üzerinden Türkiye’deki kargo firmalarına yapılmış, içerisinde “korona”, “corona”, “covid” ve “pandemi” kelimeleri geçen tüm şikâyetler bu çalışmanın veritabanını oluşturmaktadır. Veritabanında, şikâyetlerin hangi firmaya yapıldığı, içerikleri, linkleri, okunma sayıları, çözüm durumları ve tarihleri gibi çalışmada kullanılacak önemli bilgiler bulunmaktadır.

3.1.2. Web kazıma modülü

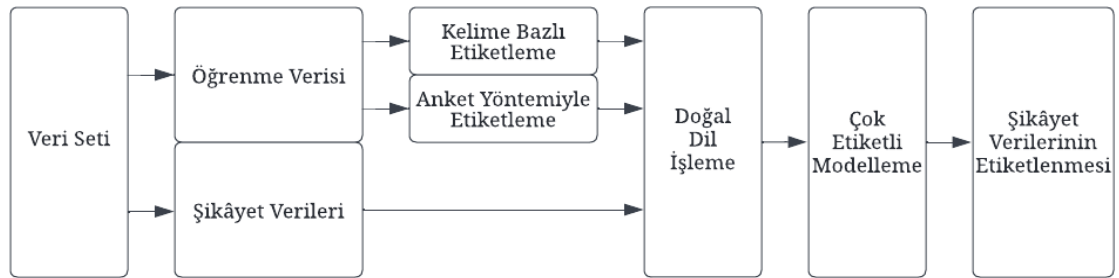
Web sitelerindeki verilerinin elde edilebilmesi için birçok farklı yöntem ve modül mevcuttur. Bunlardan biri de Python diliyle yazılmış açık kaynaklı bir Web Tarama Çerçevesi olan Scrapy modülüdür (Scrapy.org, 2021). Scrapy ile çeşitli web sitelerinden metin verileri çıkarılabilir. Çıkarılmak istenen verilerin adresleri “css selector” ya da “xpath selector” kullanılarak belirlenir ve kod içerisinde belirtilir. Çalışmada kullanılan verilerin çıkarımında bu modülden faydalanılmıştır.

3.2. Yöntem

Bu çalışmada, pandemi döneminde Türkiye’de kargo firmalarına yapılan şikâyetler makine öğrenmesi ve multilabel sınıflandırma algoritmalarından faydalanılarak sınıflandırılmış ve en çok hangi konular üzerinden bu firmaların şikâyet aldıkları tespit edilmiştir.

Literatüre ve incelenen şikâyetlerin içeriklerine göre kargo müşterilerinin en çok şikâyet ettikleri konular: “Gecikti veya Dağıtımına Çıkmadı”, “Evde yok notu düşüldü veya Kapıya Getirilmedi”, “Telefonlara Cevap Verilmedi”, “İade Süreci”, “Teslim Alınmadı veya Teslim Edilmedi” ve “Hijyen Kurallarına Uyulmadı” olarak belirlenmiş ve bu konular için 6 farklı kategori oluşturulmuştur.

Çalışmada kullanılan veri öncelikle eğitim verisi ve tüm şikâyet verisi olarak iki kısma ayrılır. Eğitim verisi, Tüm şikâyet verisinde bulunan ilk 3000 şikâyetten oluşturulur. 2. veri grubu da ilk olarak doğal dil işleme yöntemlerinden geçirilerek gürültüden arındırılmış veriler elde edilir. Elde edilen temizlenmiş verilerden eğitim verisine Python dili ile yazılmış kodlar yardımıyla ve Anket yöntemiyle katılımcılardan alınan yanıtlara göre örnek etiket atamaları yapılır. Daha sonra bu eğitim verisi multilabel algoritmalarıyla eğitilerek öğrenme gerçekleştirilir. Tüm şikâyet verisine ise eğitilen modele göre multilabel algoritmalarıyla etiket atamaları yapılarak sınıflandırma işlemi gerçekleştirilmiş olur (Şekil 3.1.).

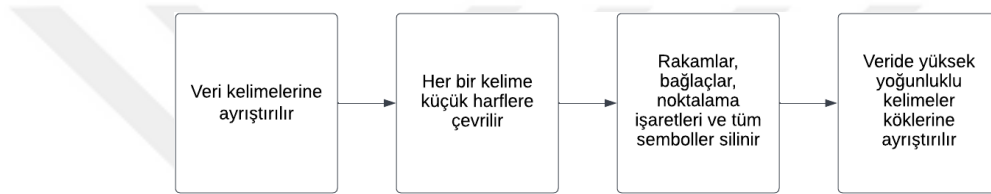


Şekil 3.1. Veri analizi süreci

Bu çalışmada uygulanan multilabel yaklaşımı içerisinde veriler NLP işlemlerinden geçirildikten sonra Etiket Ataması işlemleri gerçekleştirilmiş ve ardından Lojistik Regresyon, OneVsRestClassifier, Vektörleştirme yöntemleri kullanılmıştır.

3.2.1. Doğal dil işleme

Veritabanına kaydedilen şikâyet içerikleri, multilabel algoritmaları ile sınıflandırma işlemine başlanmadan önce ilk olarak doğal dil işleme sürecinden geçerler. Doğal dil işlemede: Tüm harfler küçük harflere çevrilir ve tüm içeriklerde yoğunluğu yüksek kelimeler köklerine ayrıştırılır (Şekil 3.2.).



Şekil 3.2. Doğal dil işleme adımları

3.2.2. Etiket ataması

Veritabanındaki şikâyet verilerinin eğitim verisi olarak ayrılan kısmında bu şikâyetlerin hangi kategorilere ait olduğunun belirlenmesi gerekmektedir. Bu sayede veriler eğitilecek ve daha sonra gelen tamamen bağımsız bir şikâyet, bu eğitim verisinden elde edilen verilere göre tahmin edilerek kategorilize edilecektir. Bu yüzden eğitim verisine etiketleme yapılması için iki yöntem belirlenmiştir. İlk yöntem kelime bazlı etiket ataması diğer yöntem ise anket yöntemi ile etiket atamasıdır.

3.2.2.1. Kelime bazlı etiket ataması

Eğitim verisi içerisinde yer alan her bir şikâyet içeriğini kendi içerisinde kelimelerine ayrıştırdıktan sonra, ilgili konu ile alakalı kelimelerin geçip geçmediğini kontrol eden ve geçiyorsa, geçtiği şikâyet içeriğini ilgili etiketlere kaydeden bir program, Python programlama dili ile yazılmıştır. İlgili etiket ile alakalı olan kelimelerden herhangi

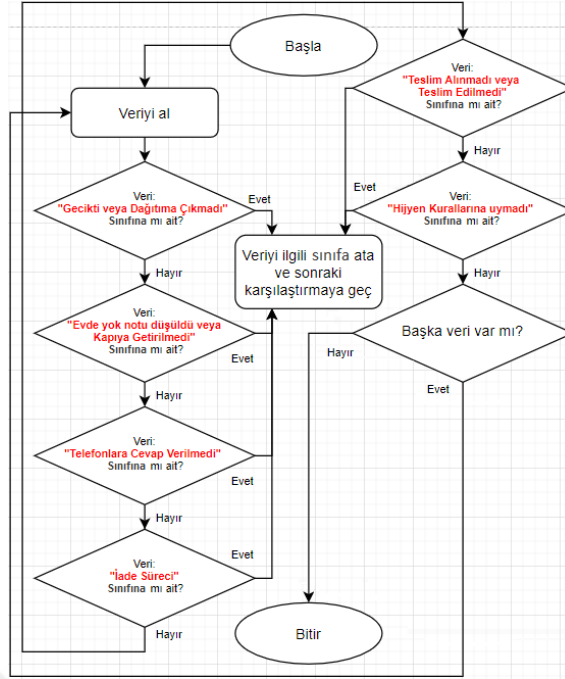
biri içerik içerisinde geçiyorsa etiket ataması işlemi program tarafından otomatik olarak gerçekleştirilmektedir (Şekil 3.3.).

Örneğin bir içerik verisinin Tablo 3.1.'deki gibi olduğu düşünüldüğünde, içerik verisi örneği doğal dil işlemeden geçtikten sonra toplam 98 kelimeden oluşmaktadır. Örnek içerisinde "maske", "iade" ve "geri gönderme" gibi programa önceden tanımlanmış spesifik kelimeler geçmektedir. Programa göre "maske", "hijyen kurallarına uyulmadı" etiketinin bir üyesi iken, "iade" ve "geri gönderme" ise "iade süreci" etiketinin bir üyesidir.

Tablo 3.1. Kargo firmalarına yapılan bir şikâyet içerik örneği

İçerik
A Kargo Covid Önlemleri Alınmıyor A Kargo Alibeyköy şubesine her gittiğimizde çalışanlarda maske olmadığını gözlemliyoruz ayrıca sıramatikleri çalışmıyor ve o havasız ortamda herkes iç içe. Sıramatik çalışmadığı halde çalışanlar bir de içeride sırasını bekleyenlere dışarıda beklemesi için bağıyor, oysa sıramatik çalışsa zaten kimse o havasız ve maskesi olmayan çalışanların olduğu ortamda beklemek istemez. İşlem sıramız geldiğinde yine maskesiz bir beyefendi bize hizmet verdi ve işlem uzadığı için maskesiz çalışanın karşısından kaçmak için yer aradık, sürekli uzak durmaya çalıştık. Çalışanların gelen kişilere ve kendilerine saygısı için maskelerini takılı tutmaları gerektiği İşletme Amiri tarafından kendilerine bildirilmelidir ki zaten İşletme amirinin de umurunda olmadığı kesin. Zaten kargom iade olmuş çoktan gerisin geri gönderilmiş, teslim alacak ürün bulamadık.

Tanımlamalara göre program tamamlandığında "iade süreci" ve "hijyen kurallarına uyulmadı" etiketleri, ilgili içeriğin alakalı olduğu içerikler olarak kabul edilerek bu etiketlere atama işlemi program tarafından gerçekleştirilmektedir.



Şekil 3.3. Kelime bazlı etiket ataması akış diyagramı

Eğitim verisi içerisinde bulunan tüm dokümanlar bu şekilde etiket ataması işleminden geçtikten sonra veri, eğitime hazır hale gelmektedir.

3.2.2.2. Anket yöntemi ile etiket ataması

Bu yöntemde amaç eğitim verisinde bulunan her bir şikâyetin farklı kişiler tarafından incelenmesi ve kategorilize edilerek etiketlere atamalar yapmalarının sağlanmasıdır. Böylelikle insan zekasından da faydalanılarak daha doğru etiketlemelerle eğitim verisinin eğitilmesi amaçlanmış olacaktır.

ŞİKAYETLER									
Dosya Düzenle Görünüm Ekle Biçim Veri Araçlar Uzmanlar Yardım Son düzenleme 13 Haziran 2021 tarihinde yapıldı									
A1									
	A	B	C	D	E	F	G	H	I
1	RANDOM SIRA	ŞİKAYET	RANDOM		İNGİLİZCESİ				
2	55	E Kargo Gülnur Korkmaz İşitme Engelliyim--> Ben işitme engelliyim karg	0.03736855195		2 E Cargo Gülnar Korkmaz deaf -> I came deaf cargo but no call was not went t				
3	851	A Kargo 10 Gün Önce Gönderilen Kargo Hâlâ Elime Ulaşmadı 7 Gündür	0.78059343365		7 A Cargo 10 days ago Posted cargo still in my hand, I reach 7 days is in distr				
4	348	C Kargo Ürünü Kapiya Getirmeme--> Merhaba geçen gün Trendyol üze	0.3136546084		8 C Cargo Products to the door to fetch -> Hi I did Trendyol day out shoppi				
5	330	B Kargo Kuryesinin Terbiyesizliği--> İnternette söylediğim paketi	0.2840391264		9 Rudeness of B Cargo Shipping Courier -> Although the courier leave the apart				
6	298	A Kargo Ofislerini Kapatılar Ve Ürünlerimizi Kaybediyoruz--> Döşemealtı	0.2845330475		14 A Office of Field Close Shipping And we're losing our products -> Döşemealtı				
7	607	C Kargo Evde Olmama Rağmen Kuryenin Evde Olmadığını İddia Edip	0.5331243478		17 C Cargo Home Absence Although the home Being that I claim Edip cargo cour				
8	17	B Kargo'da Kargom Kayıp--> 04.06.2020 tarihinde Erzurum Gürcükapı	0.01361072482		19 Lost in my B Cargo Shipping! -> Erzurum on 05.04.2020 Gürcükapı I have given t				
9	670	D Kargo Ürünü Evime Getirmeden Geri Göndermiş--> D Kargo kadar i	0.8002549021		21 D Cargo product to fetch my house back from across Europe -> D Cargo who				
10	776	A Kargo Ptt Pendik Şubesi 8 Gündür Kargo Bekletiyor--> Pendik şube sür	0.6931881715		24 A Cargo Post Office Pendik Branch 8 is day shipping put aside -> Pendik branc				
11	22	A Kargo 64 Gün Oldu Kargom Gelmiyor--> 23.07.2020 tarihinde hamburg	0.01880574701		25 A Cargo 64 Days Had not Come Shipping -> 23.07.2020 Date of hamburg tank pack				
12	73	A Kargo Kargom Getirdiklerini S--> Bunlar nasıl insan ya hem evime gellip	0.05266665028		27 A Cargo Cargo is what they brought S -> How do these people come to my house				
13	389	B Kargo'nun Sorumsuzluğu--> Ürünü dağıtma çıkıyorlar. Ücreti eve tes	0.3820251883		28 Exclusion of the B Cargo -> take out the product distribution. As I paid the				
14	869	D Kargo Adrese Gelinmeden Adreste Bulunmadınız Sorunsal--> Bir inter	0.7853309315		31 D Cargo bride before you address None Address Problematic -> shopping D C				
15	1068	D Kargo Ürün Teslim Edilmiyor--> Ümraniye çakmak Begonya şubesi kar	0.9358251527		37 D Cargo Product Delivery Not Applicable -> Ümraniye lighter Begonia branch c				

Şekil 3.4. Şikâyetleri içeren Google Sheets sayfası

Eğitim verisi içerisinde yer alan 3000 adet şikâyet öncelikli olarak Google Sheets platformuna hem Türkçe hem de İngilizce olarak aktarılmıştır (Şekil 3.4.). Böylelikle ankete erişim sağlayan yabancı kişilerin de şikâyetlerin içeriğini anlayıp anketi tamamlaması hedeflenmiştir.

Her bir şikâyet için rastgele birer sayı üretilip daha sonra bu sayılardan en büyük olanı 3000 ve en küçük olanı 1 olacak şekilde tüm şikâyetlere geçici rastgele id ler atanmaktadır. Geçici olması, bu sheets sayfasının dinamik olarak 4 dk da bir kendini yenilemesi ve her seferinde şikâyetlere atanan 1 ile 3000 arasındaki id nin farklı olmasından kaynaklanmaktadır. Bu farklılığın oluşturulmasının sebebi ise bu sheets sayfasından oluşturulan dinamik bir web sayfasında sadece 1 ile 10 arasındaki id lere sahip olan şikâyetlerin gösterilecek olmasıdır (Şekil 3.5.).

ŞİKAYETLER : KARGO				
SIRA	ŞİKAYET		COMPLAINT	NO
1	C Kargo Evde Olmama Rağmen Kargom Gelmedi-> 711176909850 numaralı gönderi ile tarafıma kargo gönderildi. Kendim şu an Covid 19 tanısı ile yaklaşık 1 haftadır evde karantina olmama rağmen güya arkadaşlar gelmiş kapıyı çalmış telefonla aramışlar ama ulaşamamışlar. Evde karantinadayım ve telefonum kesinlikle galmadı. Bu yüzden şu an kargom şubede ve kargo olarak gelende tedavim için gerekli olan ilaç. Şubeyi aradığımda da bugün getirmenin mümkün olmayacağını söylüyor. (ŞİKAYET NO: 960)	FORM	C Cargo Shipping Despite not being at home I did not come-> number 711176909850 with the shipment sent to the cargo. Despite myself to be that Covidien has approximately 1 week with 19 diagnostic quarantine at home they allegedly stole a phone call friends come across the door but could not reach. I'm definitely quarantine at home and the phone did not ring. So now required for drug treatment in branch offices and cargo from the cargo. When I call the Branch says he will not be able to bring today. (Complaint NO: 960)	1
2	A Kargo Dağıtım Çıkamayan Kargo-> 8 Temmuz 2020 tarihinde A iş yerine gelen KP0273****95 takip numaralı kargom hala dağıtım çıkmadı! Dakikalarca hatta beklememe rağmen müşteri hizmetlerine ulaşamıyorum! Pandemi nedeniyle bir çok alışverişimizi internetten yaptığımız bu dönemde A şubeye gidip paketimi teslim almak zorunda kalacağım! (ŞİKAYET NO: 1640)		A Cargo Distribution could not get out of Shipping -> Post Office came to the job on July 8, 2020 KP0273****95 cargo tracking number'm still waiting for distribution! Although I can not get to wait for minutes or even customer service! I'm going to get in this period went to deliver my package from the post office branch we did a lot of shopping our internet because the pandemic! (Complaint NO: 1640)	2
3	E Kargo Kargo Teslim Süresi-> 20.06.2020 tarihinde Ayvalık şubesine paketim gelmiş. 1 hafta olacak hala teslim etmiyorlar şubede kurye dağıtım gözükiyor 3 gündür nereden nereye dağıtıyorlar anlamıyorum. Şubeyi sitede belirtilmiş numaradan arıyoruz ağan yok. O zaman neden varlar kapatsınlar dükkanı. Ama gördüğüm kadarıyla E Kargo'dan herkes şikâyetçiyim. Bahane olarak korona sürecini gösteriyor ama bir tek Sürat Kargo'yu vurmuş anlaşılan diğer firmalar ürünleri getiriyor. Kargoyla iş çok yapıyorum ama ilk defa bir kargoyla bu kadar uğraştım (ŞİKAYET NO: 2264)		E Cargo Shipping Delivery Time -> Date 6/20/2020 package came in Ayvalık branch. 1 week still looks courier delivery will be delivered in 3 days they do not understand the branches from where they distribute. No number listed on the site are looking to open a branch. Then why would they have shut shop. But as I see E Cargo He was complaining from everyone. Show choir as an excuse, but brings the process to other companies products apparently hit a one-Surat Kargo. I'm doing a lot of work, but first a courier to courier I fought so hard (NO COMPLAINT S: 2264)	3

Şekil 3.5. Şikâyetleri içeren dinamik web sayfası

Böylelikle ilk 10 id sürekli farklı şikâyetleri işaret edeceğinden, bu web sayfasına erişen kişilerin her seferinde farklı 10 şikâyeti görmesi sağlanmıştır. Bu web sayfasına erişen kişiler sayfa üzerinde bulunan “Form” butonuna tıklayarak Google Form üzerinde oluşturulmuş olan anket sayfasına yönlendirilmektedir (Şekil 3.6.).

Anket sayfasında ise kişilerden az önce erişim sağladıkları web sayfasında bulunan 10 farklı şikâyetin, önceden belirlenmiş olan kategorilerden hangisine veya hangilerine ait olabileceklerini belirlemeleri istenir. 10 şikâyetle de etiketlemelerini yapan kişiler anketi tamamlayarak sonlandırır.

Bölüm 1/10

KARGO FİRMALARI ŞİKAYET ANKETİ / CARGO COMPANIES COMPLAINT SURVEY

Türkiye’de kargo firmalarına yapılan şikayetlerin değerlendirilmesi konusundaki ankete hoş geldiniz. Ankette sizden, bu form sayfasına erişim sağladığınız web sitesinde yer alan ve pandemi döneminde kargo firmalarına yapılan şikayetlerden rastgele 10 tanesinin hangi şikayet kategorilerine ait olduğunu belirlemeniz beklenmektedir. Her bir içerik hiçbir kategoriye uymayabileceği gibi birden fazla kategoriye de dahil olabilmektedir. Formun gidişatı, bu şikayetlerin sonunda yer alan numaralarını (şikayet no) sırası ile her bölümde formda belirtilen alana yazmanız, daha sonra içeriğe göre kategorileri belirlemeniz ve sonraki şikayete geçmeniz şeklinde ilerleyecek, son şikayetten sonra anket tamamlanacaktır.

Welcome to the survey on the evaluation of complaints made to cargo companies in Turkey. In the survey, you are expected to determine which category of complaints randomly 10 of the complaints made to cargo companies during the pandemic period on the website from which you access this form page. Each content may not fit into any category, or may be included in more than one category. The course of the form will proceed as you write the numbers (complaint no) at the end of these complaints in the field specified in the form in each section, then determine the categories according to the content and move on to the next complaint, after the last complaint the survey will be completed.

E-posta *

Geçerli e-posta

Bu form e-posta topluyor. [Ayarları değiştir](#)

Şekil 3.6. Google Form anket sayfası

Böylelikle ankete erişim sağlayan her bir kişi 10 adet şikâyeti inceleyerek etiketlemiş olur ve ne kadar çok kişi ankete erişim sağlarsa o kadar çok sayıda şikâyet verisi kategorize edilerek veri eğitime hazır hale getirilir.

3.2.3. Lojistik regresyon

Lojistik Regresyon, belirli bir sınıf veya olayın olasılığını modellemek için kullanılabilen, veriler arasında sınıflandırma işlemi yapmaya yarayan ve sayısal verilerin sınıflandırılmasında kullanılan bir "Denetimli makine öğrenimi" algoritmasıdır (Towardsdatascience.com, 2022). Sonucun sadece “1 (doğru) veya 0 (yanlış)” olarak değer alabilmesi durumunda çalışır.

3.2.4. One vs rest classifier

One-Vs-Rest Classifier yaklaşımına göre, çoklu sınıf içeren sınıflandırma problemlerinin çözümü olarak çoklu sınıfları ikili sınıflara dönüştürme yöntemi uygulanmaktadır (Machinelearningmastery.com, 2022). Yani sınıflandırma işlemi

sırasında her bir sınıf 1 değerini alırken diğer sınıflarda 0 değerini alır. Böylelikle ifade sınıflarından birinin diğerlerinden ayrıştırılması sağlanır.

3.2.5. Vektörleştirme

Vektörleştirme, girdi verilerini ham olan metin biçiminden ML modellerinin desteklediği biçim olan gerçek sayıların vektörlerine dönüştürmeye yönelik klasik bir yaklaşım için kullanılan bir terimdir. Bu yaklaşım, bilgisayarların icadından beri kullanılmakta, çeşitli alanlarda başarılı sonuçlar elde etmektedir. Günümüzde bu yaklaşım NLP'de de başarılı bir şekilde kullanılmaktadır (Neptune.ai, 2022).

3.2.5.1. Fasttext

Fasttext, Word2vec'in bir uzantısıdır ve Facebook tarafından geliştirilmiştir. Kelimeleri yapay sinir ağına tek tek girdi olarak vermek yerine birkaç harf bazlı olarak parçalar. Parçalanmış bu yapıları “n-gram” denir. Burada “n” ifadesi n-gramın kaç harften oluştuğunu gösterir. Bir kelimenin word vektörü tüm bu n-gram vektörlerinin toplamıdır (Aksu ve Karaman, 2020).

3.2.5.2. Tfidf vectorizer

TfidfVectorizer, metin verilerini vektörlere dönüştürmek ve kelimelerin vektör ölçülerini bulmak için kullanılır. TfidfVectorizer'da bir kelimenin genel belge ağırlığını dikkate alınır (Çelik ve Koç, 2021). En sık kullanılan kelimelerin tespit edilmesini sağlar. TfidfVectorizer, kelime sayılarını ve belgelerde ne sıklıkta göründüklerini ölçerek ağırlıklandırır. Metin madenciliğinde yaygın olarak kullanılan TFIDF denklemi şu şekildedir (Denklem 3.1).

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) 100 \quad (3.1)$$

Denklemde $w_{i,j}$ i kelimesinin, j dokümanındaki ağırlığını ifade etmektedir. N , veri setindeki toplam doküman sayısı, $tf_{i,j}$ i kelimesinin j dokümanında görülme sıklığı,

df_i ise i teriminin veri seti içindeki doküman sıklığı olarak ifade edilmektedir (Zhang ve ark., 2011).

3.2.5.3. Hashing vectorizer

Hashing Vectorizer, metni vektöre dönüştürmek için bir hash algoritmasını kullanır. Hash algoritması belgedeki bütün cümlelere uygulanabilmektedir (Gadde ve ark., 2021).

Hash işlevleri, rastgele uzunluktaki bir diziye sabit uzunluktaki bir diziye sıkıştırarak girdi ve özet değeri arasında benzersiz bir ilişki sağlarlar ve bu şekilde büyük miktarda bir veriyi çok daha küçük ve özgün bir değer ile değiştirirler (Deepakumara ve ark., 2001).

3.2.5.4. Count vectorizer

CountVectorizer, bir belgedeki her bir kelimenin belge içerisinde geçiş sayısını hesaplayarak bu belgeyi vektörlere dönüştüren bir yöntemdir (Tablo 3.2.) (Goyal, 2021).

Tablo 3.2. Count vectorizer örneği

kargo firmasında benim kargo kaybolmuş ve dün teslim edilmeliyken teslim edilmemiş								
kargo	firmasında	benim	kaybolmuş	ve	dün	teslim	edilmeliyken	edilmemiş
2	1	1	1	1	1	2	1	1

3.3. Analiz

Verilerin analizi sürecinde doğruluk değerlerinin tespiti için karışıklık matrisi yönteminden faydalanılmıştır.

3.3.1. Karışıklık matrisi

Karışıklık matrisi, gerçek değerlerin bilindiği bir dizi test verisi üzerindeki bir sınıflandırma modelinin (veya "sınıflandırıcının") performansını tanımlamak için sıklıkla kullanılan bir tablodur (Tablo 3.3.) (Dataschool.io, 2022).

Tablo 3.3. Karışıklık matrisi				
Yapılan toplam tahmin sayısı	Karışıklık matrisi			Toplam
	Tahmin: Yok	Tahmin: Var		
Gerçek: Yok	TN	FP		Gerçek yok
Gerçek: Var	FN	TP		Gerçek var
Toplam	Tahmin yok	Tahmin var		

Gerçek var (TP): Evet (kategoriye ait) olarak tahmin edilen ve o kategoriye ait olan durumlardır.

Gerçek yok (TN): Hayır (kategoriye ait değil) olarak tahmin edilen ve o kategoriye ait olmayan durumlardır.

Tahmin var (FP): Evet olarak tahmin edilen, ancak aslında o kategoriye ait olmayan durumlardır.

Tahmin yok (FN): Hayır tahmin edilen, ancak o kategoriye ait olan durumlardır.

Doğruluk: Genel olarak, sınıflandırıcının ne sıklıkla doğru olduğudur (Denklem 3.2).

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.2)$$

Yanlış Sınıflandırma Oranı: Genel olarak, sınıflandırıcının ne sıklıkla yanlış olduğudur. "Hata Oranı" olarak da bilinir (Denklem 3.3).

$$\text{Yanlış Sınıflandırma} = \frac{FP+FN}{TP+TN+FP+FN} \quad (3.3)$$

Gerçek Pozitif Oran: Aslında evet olduğunda, ne sıklıkla evet tahmininde bulunulduğudur. “Hassasiyet” veya “Hatırlatma” olarak da bilinir (Denklem 3.4).

$$\text{Gerçek Pozitif Oranı} = \frac{TP}{\text{Gerçek var}} \quad (3.4)$$

Yanlış Pozitif Oranı: Aslında hayır olduğunda, ne sıklıkla evet tahmininde bulunulduğudur (Denklem 3.5).

$$\text{Yanlış Pozitif Oranı} = \frac{FP}{\text{Gerçek yok}} \quad (3.5)$$

Gerçek Negatif Oran: Aslında hayır olduğunda, ne sıklıkla hayır tahmininde bulunulduğudur. “Özgüllük” olarak da bilinir (Denklem 3.6).

$$\text{Yanlış Sınıflandırma} = \frac{TN}{\text{Gerçek yok}} \quad (3.6)$$

Kesinlik: Evet tahmin edildiğinde, ne sıklıkla doğru olduğudur (Denklem 3.7).

$$\text{Kesinlik} = \frac{TP}{\text{Tahmin var}} \quad (3.7)$$

Yaygınlık: Evet koşulunun gerçekte ne sıklıkta ortaya çıktığıdır (Denklem 3.8).

$$\text{Yaygınlık} = \frac{\text{Gerçek var}}{TP+TN+FP+FN} \quad (3.8)$$

BÖLÜM 4. ARAŞTIRMA BULGULARI

İçerisinde “korona”, “corona”, “covid” ve “pandemi” kelimeleri geçen tüm şikâyetler, sikayetvar.com web sitesi üzerinden listelenmiş ve Scrapy modülü kullanılarak Python dilinde yazılan kodlar yardımıyla bu şikâyetlerin hangi firmaya yapıldığı, içerikleri, linkleri, okunma sayıları ve tarihleri gibi önemli bilgiler çekilerek veritabanına kaydedilmiştir. Daha sonra bu veritabanı içerisindeki tüm içerikler yine Python dilinde yazılan kodlar yardımıyla kelimelerine ayrıştırılarak, her kelimenin tüm içerik içerisinde toplamda kaç kez geçtiği saydırılmıştır. Elde edilen sonuçlara göre tüm içerikte “kargo” kelimesinin toplamda 27555 defa geçerek en çok kullanılan kelime olduğu tespit edilmiştir (Tablo 4.1.).

Tablo 4.1. Tüm içeriklerde en çok geçen ilk 10 kelime

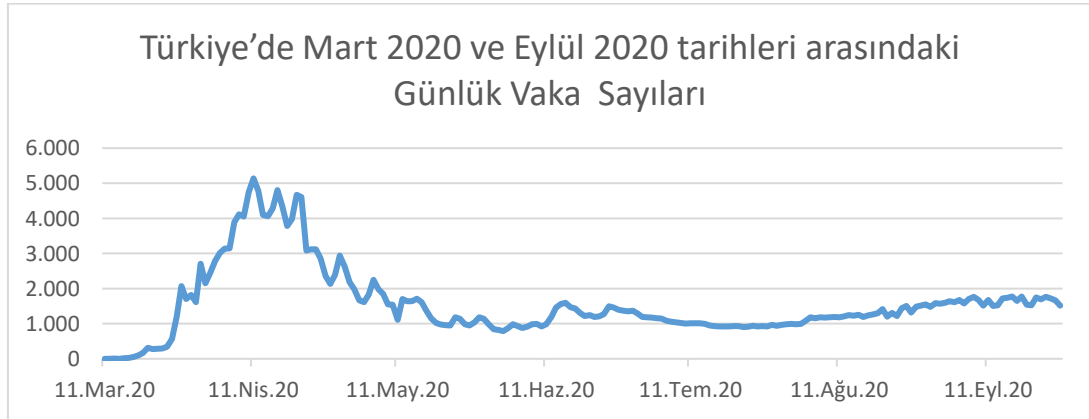
Sıra	Sıklık	Kelime
1	27555	Kargo
2	13504	Şube
3	8734	Teslim
4	7998	Ürün
5	4807	E_kargo
6	4747	Telefon
7	4444	Adres
8	4191	Pandemi
9	3719	Müşteri
10	3637	D_kargo

Daha sonra veritabanından sadece kargo firmalarına yapılan şikâyet verileri ayrıştırılmış ve farklı bir veritabanı olarak kaydedilmiştir. Yeni oluşturulan veritabanında Mart 2020 ve Eylül 2020 tarihleri arasındaki verilerden elde edilen sonuçlara göre şikâyetlerin zamana göre yoğunluk haritası çıkartılmıştır (Şekil 4.1.).



Şekil 4.1. Mart 2020 ve Eylül 2020 tarihleri arasındaki şikâyetlerin zamana göre yoğunluk haritası

Buna göre şikâyetlerin en yoğun olarak Mayıs 2020 sonu ile Temmuz 2020 başı arasında yapıldığını gözlemlemek mümkündür. Türkiye’deki günlük vaka sayıları incelendiğinde ise Şekil 4.2.’de görüldüğü gibi vakaların en yoğun olarak Nisan 2020 başı ile Mayıs 2020 sonu arasında ortaya çıktığı gözlemlenmektedir (Covid19.saglik.gov.tr, 2020).



Şekil 4.2. Türkiye’deki Mart 2020 ve Eylül 2020 tarihleri arasındaki günlük vaka sayıları

Bu sonuçlara göre; Mart 2020-Eylül 2020 arasındaki veriler incelendiğinde Nisan 2020-Mayıs 2020 arasında zirve yapan corona virüs vakalarının ardından Mayıs 2020-Temmuz 2020 arasında şikâyetlerin de yoğunlaştığı gözlemlenmiştir.

Koronavirüs pandemisi başladığı günden beri günlük vaka sayıları bazı günler oldukça yüksek olmasına rağmen bazı günlerde de düşük çıkmıştır. Mart 2020 ve Nisan 2022 tarihleri arasındaki Covid-19 vakası verileri incelendiğinde en yüksek kaydedilen günlük vaka sayısı 111.157 kişi ile 4 Şubat 2022 tarihinde kaydedilmiştir (Şekil 4.3.). 28 Nisan 2022 tarihi itibarıyla de toplam vaka sayısının 15.028.397 olduğu görülmektedir (Şekil 4.4.) (Covid19.saglik.gov.tr, 2022).



Şekil 4.3. Türkiye'deki Mart 2020 ve Nisan 2022 tarihleri arasındaki günlük vaka sayıları



Şekil 4.4. Türkiye'deki Mart 2020 ve Nisan 2022 tarihleri arasındaki toplam vaka sayısı

Şekil 4.5.'de gösterilen Mart 2020 ve Nisan 2022 tarihleri arasındaki günlük şikâyet verileri ve Şekil 4.3.'de yine aynı tarih aralığında gösterilen günlük vaka sayısı

verileri birlikte incelendiğinde; vaka sayısının arttığı dönemlerde şikâyet sayılarında da artış olduğu görülmektedir. Ancak vaka sayılarının arttığı dönemlerde zaman ilerledikçe vaka yoğunluğunun arttığı gözlemlenirken ters orantılı olarak şikâyet yoğunluğunun da azaldığı gözlemlenmiştir.



Şekil 4.5. Türkiye’de Mart 2020 ve Nisan 2022 tarihleri arasındaki günlük şikâyet sayısı

4.1. Kelime Bazlı Etiket Ataması ile Eğitilen Model

Eğitim verilerini oluşturmak için Şekil 4.6.’da gösterildiği gibi kod bloklarından oluşan ve her bir kategori için o kategoriyle ilişkili kelime ve kelime gruplarını seçerek o kategoriye dahil eden bir bilgisayar programı yazılmıştır. Oluşturulan programın hangi kelime ve kelime gruplarını seçerek kategorize ettiği Tablo 4.2.’de gösterilmiştir. Bu şekilde ilk 3000 adet şikâyet verisi kategorilize edilerek diğer şikâyetlerin tahmin edilebilmesi için bir eğitim verisi hazırlanmış olur.

```

280 if 'dağıtım' in icerik:
281     dagitimdegisken=dagitimdegisken+1
282 if 'çıkmadı' in icerik:
283     dagitimdegisken=dagitimdegisken+1
284 if 'çıkış gube' in icerik:
285     dagitimdegisken=dagitimdegisken+1
286 if 'gelmedi' in icerik:
287     dagitimdegisken=dagitimdegisken+1
288 if 'beklet' in icerik:
289     dagitimdegisken=dagitimdegisken+1
290 if 'süredir' in icerik:
291     dagitimdegisken=dagitimdegisken+1
292 if 'gecik' in icerik:
293     dagitimdegisken=dagitimdegisken+1
294 if 'bekli' in icerik:
295     dagitimdegisken=dagitimdegisken+1
296 if 'hala çıkış' in icerik:
297     dagitimdegisken=dagitimdegisken+1
298 if 'geç' in icerik:
299     dagitimdegisken=dagitimdegisken+1
300 if (dagitimdegisken*100)/len(splitWords)>oran:
301     print("Gecikti veya Dağıtıma Çıkmadı")
302     print(dagitimdegisken)
344 if 'telefonu' in icerik:
345     telefondegisken=telefondegisken+1
346 if 'telefona' in icerik:
347     telefondegisken=telefondegisken+1
348 if 'telefonlar' in icerik:
349     telefondegisken=telefondegisken+1
350 if 'açmadı' in icerik:
351     telefondegisken=telefondegisken+1
352 if 'cevap' in icerik:
353     telefondegisken=telefondegisken+1
354 if 'açmıyor' in icerik:
355     telefondegisken=telefondegisken+1
356 if 'ulaşamıyor' in icerik:
357     telefondegisken=telefondegisken+1
358 if 'ulaşama' in icerik:
359     telefondegisken=telefondegisken+1
360 if 'ağır yok' in icerik:
361     telefondegisken=telefondegisken+1
362 if (telefondegisken*100)/len(splitWords)>oran:
363     print("Telefonlara Cevap Verilmedi")
364     print(telefondegisken)

```

Şekil 4.6. Şikâyetleri kategorize eden kod bloğu örneği

Tablo 4.2. Şikâyetlerin kelime ve kelime öbeği bazlı kategorize edilmesi

Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
dağıtım	adresiniz	telefonu	geri gön	teslimat yapılma	hijyen
çıkmadı	adreste	telefona	iade	teslimatı yapılma	maske
çıkış şube	evde	telefonlar	geri yolla	teslimat yapma	temiz
gelmedi	not	açmadı	geri dön	teslim edilme	titiz
beklet	evime	cevap	geri git	teslim alınma	sosyal mesafe
süredir	kapım	açmıyor		teslim etme	
gecik	çalmıyor	ulaşlamıyor		teslim alma	
bekli	çalmadı	ulaşamı			
hala çıkış	çalmadan	açan yok			
geç	geldik				
	yoktunuz				
	kağı				
	daire				
	bina				
	aşağı				
	inmem				

4.1.1. Kelime bazlı fasttext yaklaşımı

Doğal dil işleme yöntemleriyle ve kelime bazlı etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (3000 adet eğitim verisi) wordvec özelliği olan skipgram modeline göre unsupervised (denetimsiz) olarak eğitilir. Eğitilen veriye göre "corona" kelimesinin kelime vektörleri yazdırılır (Tablo 4.3.).

Tablo 4.3. Fasttext modelde "corona" kelime vektörleri

-0,23410991	0,2749716	0,39827532	-0,00576531	0,40223876	0,35967922
0,01233825	0,26752567	-0,17067723	-0,2511951	-0,6006947	-0,07042032
-0,32890987	0,12698856	-0,05223033	0,07015675	-0,5479473	0,17232268
-0,24749927	-0,17312318	0,02599316	0,17330591	0,12005164	-0,07339692
-0,46979564	0,41532096	-0,10186651	-0,08688819	0,24925774	0,37925142
0,04206943	-0,06600217	-0,3457961	0,27770385	-0,04453028	0,04893459
0,44790328	-0,46009803	0,16494945	0,22422683	0,23901227	-0,07038168
-0,1243773	0,25672212	0,2685277	0,1474138	0,00220102	-0,52599514
-0,09132313	0,58407676	-0,10875293	0,51362383	0,1915203	-0,3377935
0,20975603	-0,30483556	-0,20219	-0,26937953	-0,25305182	-0,23895411
0,14893639	-0,09208851	-0,19263919	0,39352217	-0,1428715	0,2655304
0,29645142	-0,06190142	0,12029264	0,47512463	0,21027972	0,28202724
-0,15760067	0,27039522	-0,23851101	-0,06679559	-0,0141941	-0,00329904
0,18365921	0,01315057	0,32200745	0,30285987	-0,13797835	-0,40267423
-0,26872748	0,38222492	-0,14022785	0,48111862	0,1784723	-0,38789225
-0,27079308	0,08471849	0,06070691	-0,27021363	-0,20689873	0,12307029
-0,001348	0,0880405	0,07471974	0,02627501		

“corona”, “covid”, “araba” ve “eldiven” kelimelerine vektör bakımından en yakın kelimeler yazdırılıp sonuçların benzer anlamlı kelimeler olduğu görülür (Tablo 4.4.).

Tablo 4.4. Fasttext modelde “corona”, “covid”, “araba” ve “eldiven” kelimelerine yakın kelimeler

Corona	Covid	Araba	Eldiven
Korona	Pandemisi	Arabayı	Eldiveni
Morona	Salgının	Arabaya	Eldivensiz
Coronavirüs	Salgınına	Arabada	Masksizdi
Koronavirüs	Coronavirüs	Arabayla	Maskeli
Virüs	Salgını	Arabam	Masksiz
Virüsüne	Salgın	Arabası	Maskesi
Virüsü	Koronavirüs	Arabanın	Hijyen
Virüsün	Corona	Araçla	Maske
Virüste	Salgınında	Aracı	Maskeyi
Virüslü	Virüsüne	Aracın	Maskeyle

Sonuç olarak fasttext modeli içerikteki kelimelerin benzerlerini ve yakın anlamdaki diğer kelimeleri bulma konusunda başarılı bulunmuştur.

Fasttext modeliyle multilabel sınıflandırma işlemine başlamak için ise öncelikle kargo içerik verileri fasttext modelinin desteklediği etiket formatına göre düzenlenmiş haliyle supervised (denetimli) olarak eğitilir. Daha sonra test edilmesi amacıyla 5 adet örnek metin verisi seçilir ve fasttext modeline göre bu örnek metin verilerinin her biriyle en uyumlu 6 etiket uyum sırasına göre uyum oranları ile birlikte yazdırılır.

Tahmin işlemleri ne kadar tekrarlandıysa da ilk 6 etiket hep aynı sırada tekrar etmiş ve gerçeklikten uzak sonuçlar göstermiştir. Her tahminin bu şekilde benzer sonuçlar getirmesi modelin verileri sınıflandırma konusunda bu veriler ile başarılı bir uyum yakalayamadığını düşündürmektedir.

Sonuç olarak fasttext modeli, benzer kelimelerin tespiti konusunda birçok örnekte başarılı bir sonuç veriyse de multilabel olarak veri sınıflandırma konusundaki sonuçları olumsuz olmuştur.

Tablo 4.5. Fasttext orjinal örneği

__label__sauce __label__cheese	How much does potato starch affect a cheese sauce recipe?
__label__food-safety __label__acidity	Dangerous pathogens capable of growing in acidic environments
__label__cast-iron __label__stove	How do I cover up the white spots on my cast iron stove?
__label__restaurant	Michelin Three Star Restaurant; but if the chef is not there
__label__knife-skills __label__dicing	Without knife skills, how can I quickly and accurately dice vegetables?
__label__storage-method __label__equipment __label__bread	What's the purpose of a bread box?

Fasttext orijinal örneğinde (Tablo 4.5.) olduğu gibi kısa içerikler üzerinde etiket atamaları ve denemeler yapılmıştır. Ancak bu çalışmadaki veriler fastText multilabel kargo şikâyet örneğinde (Tablo 4.6.) olduğu gibi oldukça uzun ve karmaşık veriler olarak görülmektedir. Fasttext modelinin bu veriler üzerindeki başarısızlığının sebebinin bu şekilde uzun ve karmaşık verilerin çözümlenememesinden kaynaklanıyor olabileceği düşünülmektedir. Bu sebepten ötürü farklı bir model arayışına gidilmiştir.

Tablo 4.6. Fasttext multilabel kargo şikâyet örneği

__label__ Evde-yok-notu-düşüldü-veya-Kapıya-Getirilmedi __label__ Telefonlara-Cevap-Verilmedi E Kargo	Gönlür Korkmaz İşitme Engelliyim--> Ben işitme engelliyim kargo geldi ama hiç aramadı not bırakıp gitti biz şubeyi gitmiyoruz eşim Covid 19 var bu yüzden biz karantinadayız biz çok aradık hiç cevap vermiyor terbiyesizlik Allah ıslah etsin ikinci dağıtımı kabul etmiyor suçu ne sağlık bakanı şikâyet ettim bekliyorum. Ankara ilçe Keçiören Esentepe E Kargo şikâyet edin
__label__ Telefonlara-Cevap-Verilmedi __label__ Teslim-Alınmadı-veya-Teslim-Edilmedi B Kargo Teslimat	Yapmamak İçin Ellerinden Geleni Yapan Elemanlar--> Kargo iş yerimin adresine gelecekti Covid 19 tedbirleri kapsamında hes kodu uygulamasına geçildiği için güvenlik kargo çalışanından hes kodu istemiş fakat çalışan ben vermem diyerek beni aradı kargonuz var gelin alın dedi. Tamam dedim 2 dakikaya geliyor ama gittiğimde gitmişlerdi üstelik geri aradım açmadılar. Dilovası şubeyi aradım başka adrese yönlendirin dedim tamam dediler ama onu da yapmadılar. Şimdi ben iş yerinden izin alacağım 10 TL yol parası vereceğim kargomu almak için bu ne sorumsuzluktur. Anlamak mümkün değil
__label__ Gecikti-veya-Dağıtıma-Çıkmadı __label__ Kötü-Diyalog-Veya-Saygısız-Tutum __label__ Evde-yok-notu-düşüldü-veya-kapıya-Getirilmedi A Kargo Covid Önlemleri Alınmıyor--> A Kargo Alibeyköy	şubesine her gittiğimizde çalışanlarda maske olmadığını gözlemliyoruz ayrıca sıramatikleri çalışmıyor ve o havasız ortamda herkes iç içe. Sıramatik çalışmadığı halde çalışanlar bir de içeride sırasını bekleyenlere dışarıda beklemesi için bağırıyor, oysa sıramatik çalışsa zaten kimse o havasız ve maskesi olmayan çalışanların olduğu ortamda beklemek istemez. İşlem sıramız geldiğinde yine maskesiz bir beyefendi bize hizmet verdi ve işlem uzadığı için maskesiz çalışanın karşısından kaçmak için yer aradık, sürekli uzak durmaya çalıştık. Çalışanların gelen kişilere ve kendilerine saygısı için maskelerini takılı tutmaları gerektiği İşletme Amiri tarafından kendilerine bildirilmelidir ki zaten İşletme amirinin de umurunda olmadığı kesin. Zaten kendisi 12.30 olmadan çoktan öğle arasına çıkmıştı, şikâyet edecek yetkili bulamadık.

4.1.2. Kelime bazlı tfidf vectorizer yaklaşımı

FastText modelinin, bu çalışmanın veritabanı üzerinde başarısız bir multilabel sınıflandırma işlemi yapmasının ardından farklı bir kelime vektörleme yaklaşımı olarak Tf-Idf Vectorizer yaklaşımı ile denemeler yapılmıştır.

Doğal dil işleme yöntemleriyle ve kelime bazlı etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (3000 adet eğitim verisi) tahminleme modeli olarak “LogisticRegression”, sınıflandırma yöntemi olarak “OneVsRestClassifier” ve vektörleme yaklaşımı olarak da “TfidfVectorizer” modeli yardımıyla eğitilir. Eğitilen veri ile oluşturulmuş modelin test edilmesi amacıyla 10 adet örnek metin verisi seçilir ve sonuçlara göre bu örnek metin verilerinin her biriyle en uyumlu etiketler uyum oranları ile birlikte yazdırılır.

10 örnek metin verisi üzerinde gerçekleştirilen tahmin işlemlerinin makul sonuçlar verdiği gözlemlendiğinden dolayı bir diğer test için, eğitimde kullanılan 3000 adet eğitim verisi üzerinde bu model uygulanarak sonuçlar izlenmiştir (Tablo 4.7.).

Tablo 4.7. 3000 eğitim verisinde ilk 10 örnek üzerinde TfidfVectorizer tahmin oranları

Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0,2612	0,4362	0,44955	0,1009	0,0873	0,0633
0,3179	0,2796	0,4277	0,2112	0,2756	0,094
0,6167	0,3059	0,26176	0,1664	0,1481	0,5855
0,3715	0,1817	0,64245	0,0859	0,1601	0,0408
0,6186	0,2144	0,16975	0,0903	0,0893	0,0556
0,8192	0,0525	0,63114	0,0515	0,0711	0,0454
0,366	0,5344	0,39272	0,0934	0,0934	0,0618
0,1934	0,5776	0,21424	0,0985	0,145	0,0979
0,6737	0,3957	0,32818	0,1332	0,2315	0,1978
0,4773	0,8343	0,11062	0,1045	0,0656	0,0588

Tablo 4.8. 3000 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak TfIdfVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0	1	1	0	0	0
0	0	1	1	1	1
1	0	0	0	0	1
0	0	1	0	0	0
1	0	0	0	0	0
1	0	1	0	0	0
0	1	0	0	0	0
0	1	0	0	0	1
1	0	0	0	0	1
0	1	0	0	0	0

Tablo 4.8.'de gösterilen tablo, Tablo 4.7.'de gösterilen tablodaki değerlerin 1 ve 0'a dönüştürülmüş halidir. Tablo 4.7.'de bulunan her bir kategori içerisindeki değerlerin ortalaması bulunmuş ve ortalamanın altında kalan değerlerin 0'a üstünde kalan değerlerin 1'e dönüştürüldüğü ayrı bir tablo oluşturulmuştur. Bu tablo her bir şikâyetin o kategoriye ait olup olmaması ihtimalini net olarak sunmaktadır.

```

C:\Windows\System32\cmd.exe
max 2999.000000 1.000000 ... 1.000000 1.000000
0
[8 rows x 7 columns]
0 Kargo Gülnur Korkmaz İşitme Engelliyim---
1 Kargo Teslimat Yapmamak İçin Ellerinden Ge...
2 Kargo Covid Önlemleri Alınmıyor--> PTT Kar...
3 Kargo Ürünü Adrese Getirmeme, Müşteri Hizm...
4 Kargo Kargom Yola Çıkmıyor--> Şubeye gitt...
Name: içerik, dtype: object
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2999 entries, 0 to 2998
Data columns (total 8 columns):
# Column Non-Null Count Dtype
---
0 id 2999 non-null int64
1 içerik 2999 non-null object
2 Gecikti veya Dağıtıma Çıkmadı 2999 non-null int64
3 Evde yok notu düşüldü veya Kapıya Getirilmedi 2999 non-null int64
4 Telefonlara Cevap Verilmedi 2999 non-null int64
5 İade Süreci 2999 non-null int64
6 Teslim Alınmadı veya Teslim Edilmedi 2999 non-null int64
7 Hijyen Kurallarına Uyulmadı 2999 non-null int64
dtypes: int64(7), object(1)
memory usage: 175.8+ KB

TfidfVectorizer yöntemi ile train edilen 2999 test verisinde yapılan toplam 17994 tahminden 16400 kadarı doğrudur
doğruluk oranı= %91.14149160831388
d:\Users\TKPC\Desktop\multilabel kelime>

```

Şekil 4.7. TfIdfVectorizer ile tahminlenen verilerin, kelime bazlı etiket atamalarıyla arasındaki benzerlik oranı

Yapılan karşılaştırmalar sonucu 3000 eğitim verisi üzerinde yapılan tahminleme işlemi sonuçlarının, tahminlenen verilerdeki gerçek etiket atamaları ile arasında

ortaya çıkan benzerlik oranı %91,14 olarak hesaplanmıştır. Tahminleme işlemi süresince 17994 tahmin yapılmış ve 16400 tanesi doğru tahmin edilmiştir. Bu oldukça yüksek ve başarılı bir sonuçtur (Şekil 4.7.).

Elde edilen bu başarılı test sonuçlarının ardından 16333 adet olan toplam şikâyet içerikleri üzerinde bu modelin uygulamasına geçilmiştir. Uygulama sonuçlarına göre: “Gecikti veya Dağıtıma Çıkmadı” konusunda 8156, “Evde yok notu düşüldü veya Kapiya Getirilmedi” konusunda 6176, “Telefonlara Cevap Verilmedi” konusunda 6902, “İade Süreci” konusunda 4189, “Teslim Alınmadı veya Teslim Edilmedi” konusunda 5696 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış 3104 şikâyet tespit edilmiştir (Tablo 4.9.).

Tablo 4.9. Tfidfvectorizer ile kelime bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı

Kategori	Örnek Sayısı
Gecikti veya Dağıtıma Çıkmadı	8156
Evde yok notu düşüldü veya Kapiya Getirilmedi	6176
Telefonlara Cevap Verilmedi	6902
İade Süreci	4189
Teslim Alınmadı veya Teslim Edilmedi	5696
Hijyen Kurallarına Uyulmadı	3104

4.1.3. Kelime bazlı count vectorizer yaklaşımı

Doğal dil işleme yöntemleriyle ve kelime bazlı etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (3000 adet eğitim verisi) tahminleme modeli olarak “LogisticRegression”, sınıflandırma yöntemi olarak “OneVsRestClassifier” ve vektörleme yaklaşımı olarak da “CountVectorizer” modeli yardımıyla eğitilir. Eğitilen veri ile oluşturulmuş modelin test edilmesi amacıyla 10 adet örnek metin verisi seçilir ve sonuçlara göre bu örnek metin verilerinin her biriyle en uyumlu etiketler uyum oranları ile birlikte yazdırılır (Tablo 4.10.).

Tablo 4.10. 3000 eğitim verisinde ilk 10 örnek üzerinde CountVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0,0072	0,9237	0,8929	0,0047	0,0185	0,0085
0,0173	0,0285	0,8840	0,0205	0,8197	0,0277
0,9496	0,0571	0,0192	0,0317	0,0082	1,0000
0,0581	0,0112	0,9952	0,0019	0,0216	0,0007
0,9189	0,0279	0,0095	0,0022	0,0042	0,0043
0,9810	0,0007	0,9446	0,0060	0,0063	0,0145
0,1455	0,7743	0,3371	0,0035	0,0055	0,0093
0,0076	0,9593	0,0049	0,0062	0,0331	0,0258
0,9361	0,1000	0,0468	0,0109	0,0425	0,9351
0,2590	0,9999	0,0065	0,0112	0,0009	0,0024

Tablo 4.11.'de gösterilen tablo, Tablo 4.10.'da gösterilen tablodaki değerlerin 1 ve 0'a dönüştürülmüş halidir.

Tablo 4.11. 3000 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak CountVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0	1	1	0	0	0
0	0	1	0	1	0
1	0	0	0	0	1
0	0	1	0	0	0
1	0	0	0	0	0
1	0	1	0	0	0
0	1	0	0	0	0
0	1	0	0	0	0
1	0	0	0	0	1
0	1	0	0	0	0

```

C:\Windows\System32\cmd.exe
extra_warning_msg= LOGISTIC_SOLVER_CONVERGENCE_MSG)
C:\Users\TKPC\AppData\Local\Programs\Python\Python37-32\lib\site-packages\sklearn\linear_model\_logistic.py:764: Converge
enceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
  https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
  https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
extra_warning_msg= LOGISTIC_SOLVER_CONVERGENCE_MSG)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2999 entries, 0 to 2998
Data columns (total 8 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   id                                    2999 non-null   int64
1   icerik                               2999 non-null   object
2   Gecikti veya Dağıtıma Çıkmadı        2999 non-null   int64
3   Evde yok notu düşüldü veya Kapıya Getirilmedi  2999 non-null   int64
4   Telefonlara Cevap Verilmedi         2999 non-null   int64
5   İade Süreci                          2999 non-null   int64
6   Teslim Alınmadı veya Teslim Edilmedi  2999 non-null   int64
7   Hijyen Kurallarına Uyulmadı         2999 non-null   int64
dtypes: int64(7), object(1)
memory usage: 175.8+ KB

CountVectorizer yöntemi ile train edilen 2999 test verisinde yapılan toplam 17994 tahminden 17965 kadarı doğrudur
doğruluk oranı= %99.83883516727798

d:\Users\TKPC\Desktop\multilabel kelime>

```

Şekil 4.8. Countvectorizer ile tahminlenen verilerin, kelime bazlı etiket atamalarıyla arasındaki benzerlik oranı

Yapılan karşılaştırmalar sonucu 3000 eğitim verisi üzerinde yapılan tahminleme işlemi sonuçlarının, tahminlenen verilerdeki gerçek etiket atamaları ile arasında ortaya çıkan benzerlik oranı %99,84 olarak hesaplanmıştır. Tahminleme işlemi süresince 17994 tahmin yapılmış ve 17965 tanesi doğru tahmin edilmiştir. Bu oldukça yüksek ve başarılı bir sonuçtur (Şekil 4.8.).

Elde edilen bu başarılı test sonuçlarının ardından 16333 adet olan toplam şikâyet içerikleri üzerinde bu modelin uygulamasına geçilmiştir. Uygulama sonuçlarına göre: “Gecikti veya Dağıtıma Çıkmadı” konusunda 8614, “Evde yok notu düşüldü veya Kapıya Getirilmedi” konusunda 6336, “Telefonlara Cevap Verilmedi” konusunda 7610, “İade Süreci” konusunda 3431, “Teslim Alınmadı veya Teslim Edilmedi” konusunda 5100 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış 1078 şikâyet tespit edilmiştir (Tablo 4.12.).

Tablo 4.12. CountVectorizer ile kelime bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	
Kategori	Örnek Sayısı
Gecikti veya Dağıtıma Çıkmadı	8614
Evde yok notu düşüldü veya Kapıya Getirilmedi	6336
Telefonlara Cevap Verilmedi	7610
İade Süreci	3431
Teslim Alınmadı veya Teslim Edilmedi	5100
Hijyen Kurallarına Uyulmadı	1078

4.1.4. Kelime bazlı hashing vectorizer yaklaşımı

Doğal dil işleme yöntemleriyle ve kelime bazlı etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (3000 adet eğitim verisi) tahminleme modeli olarak “LogisticRegression”, sınıflandırma yöntemi olarak “OneVsRestClassifier” ve vektörleme yaklaşımı olarak da “HashingVectorizer” modeli yardımıyla eğitilir. Eğitilen veri ile oluşturulmuş modelin test edilmesi amacıyla 10 adet örnek metin verisi seçilir ve sonuçlara göre bu örnek metin verilerinin her biriyle en uyumlu etiketler uyum oranları ile birlikte yazdırılır (Tablo 4.13.).

Tablo 4.13. 3000 eğitim verisinde ilk 10 örnek üzerinde HashingVectorizer tahmin oranları

Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0,1975	0,3536	0,4399	0,0620	0,0491	0,0587
0,3540	0,3037	0,3534	0,2365	0,1558	0,0953
0,5368	0,3132	0,3024	0,1879	0,1399	0,6123
0,4074	0,1560	0,5900	0,0882	0,1403	0,0439
0,5535	0,2036	0,1692	0,0713	0,0602	0,0530
0,7947	0,0474	0,5910	0,0450	0,0441	0,0467
0,3889	0,4591	0,3812	0,0821	0,0519	0,0622
0,1392	0,5084	0,2278	0,0804	0,0849	0,1422
0,6436	0,4600	0,3690	0,1232	0,2529	0,1953
0,4720	0,7712	0,0834	0,0913	0,0391	0,0546

Tablo 4.14.’de gösterilen tablo, Tablo 4.13.’te gösterilen tablodaki değerlerin 1 ve 0’a dönüştürülmüş halidir.

Tablo 4.14. 3000 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak HashingVectorizer tahmin oranları

Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapiya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0	0	1	0	0	0
0	0	0	1	0	1
0	0	0	0	0	1
0	0	1	0	0	0
0	0	0	0	0	0
1	0	1	0	0	0
0	1	0	0	0	0
0	1	0	0	0	1
1	1	0	0	1	1
0	1	0	0	0	0

```

C:\Windows\System32\cmd.exe
000000 1.000000
[8 rows x 7 columns]
0 Kargo Gülnur Korkmaz İşitme Engelliyim----
1 Kargo Teslimat Yapmamak İçin Ellerinden Ge...
2 Kargo Covid Önlemleri Alınmıyor--> PTT Kar...
3 Kargo Ürünü Adrese Getirmeme, Müşteri Hizm...
4 Kargo Kargom Yola Çıkıyor--> Şubeye gitt...
Name: icerik, dtype: object
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2999 entries, 0 to 2998
Data columns (total 8 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   id                                    2999 non-null   int64
1   icerik                               2999 non-null   object
2   Gecikti veya Dağıtım Çıkmadı         2999 non-null   int64
3   Evde yok notu düşüldü veya Kapiya Getirilmedi  2999 non-null   int64
4   Telefonlara Cevap Verilmedi         2999 non-null   int64
5   İade Süreci                          2999 non-null   int64
6   Teslim Alınmadı veya Teslim Edilmedi  2999 non-null   int64
7   Hijyen Kurallarına Uyulmadı         2999 non-null   int64
dtypes: int64(7), object(1)
memory usage: 175.8+ KB

HashingVectorizer yöntemi ile train edilen 2999 test verisinde yapılan toplam 17994 tahmin
15337 kadarı doğrudur
doğruluk oranı= %85.23396687784818
d:\Users\TKPC\Desktop\multilabel kelime>

```

Şekil 4.9. HashingVectorizer ile tahminlenen verilerin, kelime bazlı etiket atamalarıyla arasındaki benzerlik oranı

Yapılan karşılaştırmalar sonucu 3000 eğitim verisi üzerinde yapılan tahminleme işlemi sonuçlarının, tahminlenen verilerdeki gerçek etiket atamaları ile arasında ortaya çıkan benzerlik oranı %85,23 olarak hesaplanmıştır. Tahminleme işlemi süresince 17994 tahmin yapılmış ve 15337 tanesi doğru tahmin edilmiştir. Bu yüksek ve başarılı bir sonuçtur (Şekil 4.9.).

Elde edilen bu başarılı test sonuçlarının ardından 16333 adet olan toplam şikâyet içerikleri üzerinde bu modelin uygulamasına geçilmiştir. Uygulama sonuçlarına göre:

“Gecikti veya Dağıtıma Çıkmadı” konusunda 8075, “Evde yok notu düşüldü veya Kapiya Getirilmedi” konusunda 5770, “Telefonlara Cevap Verilmedi” konusunda 7001, “İade Süreci” konusunda 3894, “Teslim Alınmadı veya Teslim Edilmedi” konusunda 5375 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış 4017 şikâyet tespit edilmiştir (Tablo 4.15.).

Tablo 4.15. HashingVectorizer ile kelime bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı	
Kategori	Örnek Sayısı
Gecikti veya Dağıtıma Çıkmadı	8075
Evde yok notu düşüldü veya Kapiya Getirilmedi	5770
Telefonlara Cevap Verilmedi	7001
İade Süreci	3894
Teslim Alınmadı veya Teslim Edilmedi	5375
Hijyen Kurallarına Uyulmadı	4017

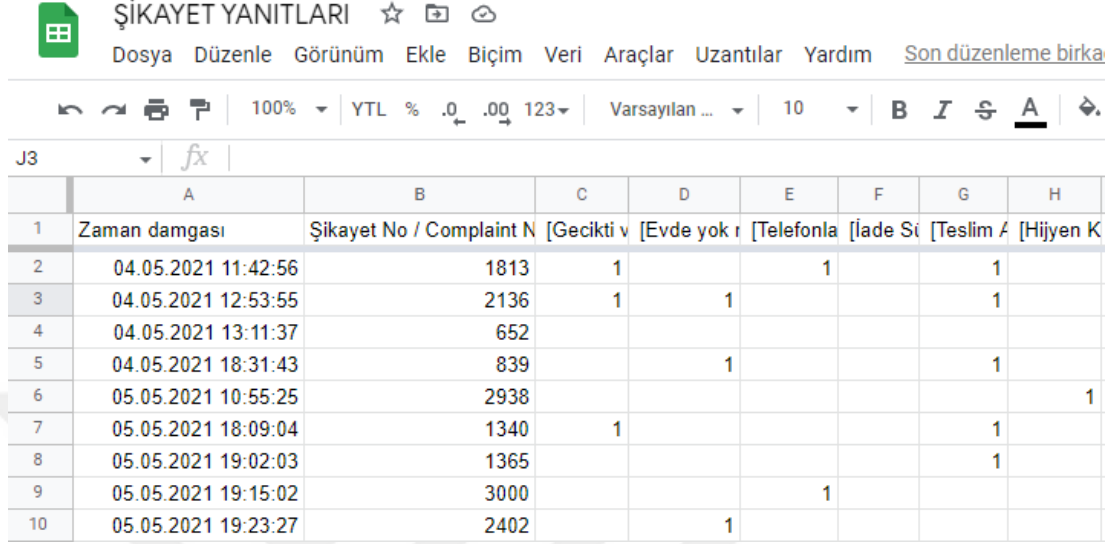
Kelime bazlı etiket ataması ile eğitilen modele göre 3000 eğitim verisi üzerinde yapılan toplamda 53982 tahmine göre hazırlanan karışıklık matrisi Tablo 4.16.’da gösterilmiştir. Yöntemin doğruluk oranı %92 olarak hesaplanmıştır.

Tablo 4.16. Kelime bazlı etiket ataması karışıklık matrisi			
Yapılan toplam tahmin sayısı = 53982	Kelime bazlı etiket ataması karışıklık matrisi		
	Tahmin: Yok	Tahmin: Var	Toplam
Gerçek: Yok	33528	1406	34934
Gerçek: Var	2874	16174	19048
Toplam	36402	17580	
Doğruluk	0,920714		

4.2. Anket Yöntemiyle Etiket Ataması ile Eğitilen Model

İnsan zekasından da faydalanılarak eğitim verisinin oluşturması için hazırlanan ve sunulan online ankete 249 kişi katılım sağlamıştır. Her bir katılımcı 10’ar adet şikâyeti sınıflandırdığından toplam 2490 sonuç elde edilmiştir. Tespit edilen ortak ve geçersiz veriler ayıklanmış, katılımcılara anket üzerinden sunulan toplam 3000 şikâyet verisinden 1887 tanesinin birbirinden farklı, geçerli ve kullanılabilir olduğu

belirlenmiştir. Şekil 4.10.'da katılımcılardan alınan yanıtlardan ilk 10 tanesi örnek olarak gösterilmektedir.



	A	B	C	D	E	F	G	H
1	Zaman damgası	Şikayet No / Complaint N	Gecikti v	Evde yok t	Telefonla	İade S	Teslim A	Hijyen K
2	04.05.2021 11:42:56	1813	1		1		1	
3	04.05.2021 12:53:55	2136	1	1			1	
4	04.05.2021 13:11:37	652						
5	04.05.2021 18:31:43	839		1			1	
6	05.05.2021 10:55:25	2938						1
7	05.05.2021 18:09:04	1340	1				1	
8	05.05.2021 19:02:03	1365					1	
9	05.05.2021 19:15:02	3000			1			
10	05.05.2021 19:23:27	2402		1				

Şekil 4.10. Anket yöntemiyle katılımcılardan alınan şikâyet yanıtları

4.2.1. Anket yöntemiyle tfidf vectorizer yaklaşımı

Doğal dil işleme yöntemleriyle ve anket yöntemiyle etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (1887 adet eğitim verisi) tahminleme modeli olarak “LogisticRegression”, sınıflandırma yöntemi olarak “OneVsRestClassifier” ve vektörleme yaklaşımı olarak da “TfidfVectorizer” modeli yardımıyla eğitilir. Eğitilen veri ile oluşturulmuş modelin test edilmesi amacıyla 10 adet örnek metin verisi seçilir ve sonuçlara göre bu örnek metin verilerinin her biriyle en uyumlu etiketler uyum oranları ile birlikte yazdırılır.

10 örnek metin verisi üzerinde gerçekleştirilen tahmin işlemlerinin makul sonuçlar verdiği gözlemlendiğinden dolayı bir diğer test için, eğitimde kullanılan 1887 adet eğitim verisi üzerinde bu model uygulanarak sonuçlar izlenmiştir (Tablo 4.17.).

Tablo 4.17. 1887 eğitim verisinde ilk 10 örnek üzerinde TfIdfVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0,3694	0,2029	0,1126	0,0535	0,2730	0,0900
0,3980	0,3385	0,2048	0,0948	0,4928	0,0959
0,1949	0,1604	0,1225	0,0683	0,2007	0,4253
0,5019	0,2197	0,2432	0,0518	0,4961	0,0670
0,5803	0,1896	0,1716	0,0640	0,3393	0,0612
0,2367	0,2286	0,1462	0,0581	0,3296	0,2413
0,2464	0,5519	0,0979	0,0482	0,3719	0,1182
0,6596	0,1889	0,3058	0,0932	0,4171	0,0664
0,4285	0,5983	0,2357	0,0949	0,3898	0,0539
0,3481	0,3671	0,0900	0,0596	0,5465	0,0700

Tablo 4.18.'de gösterilen tablo, Tablo 4.17.'de gösterilen tablodaki değerlerin 1 ve 0'a dönüştürülmüş halidir. Tablo 4.17.'de bulunan her bir kategori içerisindeki değerlerin ortalaması bulunmuş ve ortalamanın altında kalan değerlerin 0'a üstünde kalan değerlerin 1'e dönüştürüldüğü ayrı bir tablo oluşturulmuştur. Bu tablo her bir şikâyetin o kategoriye ait olup olmaması ihtimalini net olarak sunmaktadır.

Tablo 4.18. 1887 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak TfIdfVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0	0	0	0	0	1
0	1	0	1	1	1
0	0	0	0	0	1
1	0	1	0	1	0
1	0	0	0	0	0
0	0	0	0	0	1
0	1	0	0	0	1
1	0	1	1	0	0
1	1	1	1	0	0
0	1	0	0	1	0

```

C:\Windows\System32\cmd.exe
max 3000.000000 ... 1.000000

[8 rows x 7 columns]
0 Kargo Gülnur Korkmaz İşitme Engelliyim---...
1 Kargo Teslimat Yapmamak İçin Ellerinden Ge...
2 Kargo Covid Önlemleri Alınmıyor--> PTT Kar...
3 Kargo Ürünü Adrese Getirmeme, Müşteri Hizm...
4 Kargo Kargom Yola Çıkıyor--> Şubeye gitt...
Name: icerik, dtype: object
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1887 entries, 0 to 1886
Data columns (total 8 columns):
# Column Non-Null Count Dtype
---
0 id 1887 non-null int64
1 icerik 1887 non-null object
2 Gecikti veya Dağıtıma Çıkmadı 1887 non-null int64
3 Evde yok notu düşüldü veya Kapıya Getirilmedi 1887 non-null int64
4 Telefonlara Cevap Verilmedi 1887 non-null int64
5 İade Süreci 1887 non-null int64
6 Teslim Alınmadı veya Teslim Edilmedi 1887 non-null int64
7 Hijyen Kurallarına Uyulmadı 1887 non-null int64
dtypes: int64(7), object(1)
memory usage: 110.6+ KB

TfidfVectorizer yöntemi ile train edilen 1887 test verisinde yapılan toplam 11322 tahmi
nden 9603 kadarı doğrudur
doğruluk oranı= %84.81717011128775

d:\Users\TKPC\Desktop\multilabel kelime>

```

Şekil 4.11. Tfidfvectorizer ile tahminlenen verilerin, anket bazlı etiket atamalarıyla arasındaki benzerlik oranı

Yapılan karşılaştırmalar sonucu 1887 eğitim verisi üzerinde yapılan tahminleme işlemi sonuçlarının, tahminlenen verilerdeki gerçek etiket atamaları ile arasında ortaya çıkan benzerlik oranı %84,82 olarak hesaplanmıştır. Tahminleme işlemi süresince 11322 tahmin yapılmış ve 9603 tanesi doğru tahmin edilmiştir. Bu oldukça yüksek ve başarılı bir sonuçtur (Şekil 4.11.).

Elde edilen bu başarılı test sonuçlarının ardından 16333 adet olan toplam şikâyet içerikleri üzerinde bu modelin uygulamasına geçilmiştir. Uygulama sonuçlarına göre: “Gecikti veya Dağıtıma Çıkmadı” konusunda 7019, “Evde yok notu düşüldü veya Kapıya Getirilmedi” konusunda 5991, “Telefonlara Cevap Verilmedi” konusunda 6514, “İade Süreci” konusunda 4751, “Teslim Alınmadı veya Teslim Edilmedi” konusunda 7706 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış 5196 şikâyet tespit edilmiştir (Tablo 4.19.).

Tablo 4.19. Tfidfvectorizer ile anket bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı

Kategori	Örnek Sayısı
Gecikti veya Dağıtıma Çıkmadı	7019
Evde yok notu düşüldü veya Kapıya Getirilmedi	5991
Telefonlara Cevap Verilmedi	6514
İade Süreci	4751
Teslim Alınmadı veya Teslim Edilmedi	7706
Hijyen Kurallarına Uyulmadı	5196

4.2.2. Anket yöntemiyle count vectorizer yaklaşımı

Doğal dil işleme yöntemleriyle ve anket yöntemiyle etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (1887 adet eğitim verisi) tahminleme modeli olarak “LogisticRegression”, sınıflandırma yöntemi olarak “OneVsRestClassifier” ve vektörleme yaklaşımı olarak da “CountVectorizer” modeli yardımıyla eğitilir. Eğitilen veri ile oluşturulmuş modelin test edilmesi amacıyla 10 adet örnek metin verisi seçilir ve sonuçlara göre bu örnek metin verilerinin her biriyle en uyumlu etiketler uyum oranları ile birlikte yazdırılır (Tablo 4.20.).

Tablo 4.20. 1887 eğitim verisinde ilk 10 örnek üzerinde CountVectorizer tahmin oranları

Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0,1866	0,0339	0,0017	0,0011	0,0317	0,0257
0,9029	0,8538	0,7971	0,0420	0,8721	0,0146
0,0328	0,0033	0,0055	0,0107	0,0087	0,9898
0,9571	0,0812	0,0920	0,0014	0,8506	0,0035
0,9391	0,0450	0,0492	0,0101	0,0571	0,0024
0,0101	0,0427	0,0039	0,0010	0,0669	0,9675
0,0441	0,8437	0,0255	0,0058	0,1885	0,0685
0,9760	0,0360	0,1385	0,0401	0,1211	0,0099
0,1943	0,9691	0,0850	0,0178	0,0620	0,0008
0,1553	0,7832	0,0028	0,0076	0,8234	0,0064

Tablo 4.21.’de gösterilen tablo, Tablo 4.20.’de gösterilen tablodaki değerlerin 1 ve 0’a dönüştürülmüş halidir.

Tablo 4.21. 1887 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak CountVectorizer tahmin oranları

Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapiya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0	0	0	0	0	0
1	1	1	0	1	0
0	0	0	0	0	1
1	0	0	0	1	0
1	0	0	0	0	0
0	0	0	0	0	1
0	1	0	0	0	0
1	0	0	0	0	0
0	1	0	0	0	0
0	1	0	0	1	0

```

C:\Windows\System32\cmd.exe
C:\Users\TKPC\AppData\Local\Programs\Python\Python37-32\lib\site-packages\sklearn\linear_model\_logistic.py:764: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
    extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1887 entries, 0 to 1886
Data columns (total 8 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   id                                    1887 non-null   int64
1   icerik                               1887 non-null   object
2   Gecikti veya Dağıtım Çıkmadı         1887 non-null   int64
3   Evde yok notu düşüldü veya Kapiya Getirilmedi 1887 non-null   int64
4   Telefonlara Cevap Verilmedi         1887 non-null   int64
5   İade Süreci                         1887 non-null   int64
6   Teslim Alınmadı veya Teslim Edilmedi 1887 non-null   int64
7   Hijyen Kurallarına Uyulmadı         1887 non-null   int64
dtypes: int64(7), object(1)
memory usage: 110.6+ KB

CountVectorizer yöntemi ile train edilen 1887 test verisinde yapılan toplam 11322 tahmin
nden 11247 kadari doğrudur
doğruluk oranı= %99.33757286698463
d:\Users\TKPC\Desktop\multilabel kelime>

```

Şekil 4.12. Countvectorizer ile tahminlenen verilerin, anket bazlı etiket atamalarıyla arasındaki benzerlik oranı

Yapılan karşılaştırmalar sonucu 1887 eğitim verisi üzerinde yapılan tahminleme işlemi sonuçlarının, tahminlenen verilerdeki gerçek etiket atamaları ile arasında ortaya çıkan benzerlik oranı %99,34 olarak hesaplanmıştır. Tahminleme işlemi süresince 11322 tahmin yapılmış ve 11247 tanesi doğru tahmin edilmiştir. Bu oldukça yüksek ve başarılı bir sonuçtur (Şekil 4.12.).

Elde edilen bu başarılı test sonuçlarının ardından 16333 adet olan toplam şikâyet içerikleri üzerinde bu modelin uygulamasına geçilmiştir. Uygulama sonuçlarına göre:

“Gecikti veya Dağıtıma Çıkmadı” konusunda 5941, “Evde yok notu düşüldü veya Kapiya Getirilmedi” konusunda 6265, “Telefonlara Cevap Verilmedi” konusunda 5418, “İade Süreci” konusunda 2542, “Teslim Alınmadı veya Teslim Edilmedi” konusunda 6568 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış 2958 şikâyet tespit edilmiştir (Tablo 4.22.).

Tablo 4.22. CountVectorizer ile anket bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı

Kategori	Örnek Sayısı
Gecikti veya Dağıtıma Çıkmadı	5941
Evde yok notu düşüldü veya Kapiya Getirilmedi	6265
Telefonlara Cevap Verilmedi	5418
İade Süreci	2542
Teslim Alınmadı veya Teslim Edilmedi	6568
Hijyen Kurallarına Uyulmadı	2958

4.2.3. Anket yöntemiyle hashing vectorizer yaklaşımı

Doğal dil işleme yöntemleriyle ve anket yöntemiyle etiket ataması işlemleriyle işlenen kargo içerik eğitim verileri (1887 adet eğitim verisi) tahminleme modeli olarak “LogisticRegression”, sınıflandırma yöntemi olarak “OneVsRestClassifier” ve vektörleme yaklaşımı olarak da “HashingVectorizer” modeli yardımıyla eğitilir. Eğitilen veri ile oluşturulmuş modelin test edilmesi amacıyla 10 adet örnek metin verisi seçilir ve sonuçlara göre bu örnek metin verilerinin her biriyle en uyumlu etiketler uyum oranları ile birlikte yazdırılır (Tablo 4.23.).

Tablo 4.23. 1887 eğitim verisinde ilk 10 örnek üzerinde HashingVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0,3793	0,1833	0,0941	0,0413	0,2482	0,0702
0,3010	0,2536	0,1431	0,1276	0,4372	0,0948
0,1484	0,1491	0,1388	0,1028	0,1641	0,4757
0,4145	0,2038	0,2644	0,0567	0,4608	0,0813
0,5617	0,1600	0,2323	0,0622	0,3554	0,0513
0,2194	0,2226	0,1585	0,0537	0,3618	0,2655
0,2214	0,5030	0,0880	0,0564	0,3441	0,1161
0,6237	0,2077	0,3436	0,1127	0,4961	0,0506
0,4135	0,5424	0,3018	0,1028	0,4164	0,0712
0,3280	0,3390	0,0767	0,0544	0,4786	0,0656

Tablo 4.24.'de gösterilen tablo, Tablo 4.23.'de gösterilen tablodaki değerlerin 1 ve 0'a dönüştürülmüş halidir.

Tablo 4.24. 1887 eğitim verisinde ilk 10 örnek üzerinde 1 ve 0 olarak HashingVectorizer tahmin oranları

Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
0	0	0	0	0	0
0	0	0	1	0	1
0	0	0	1	0	1
0	0	1	0	1	1
1	0	1	0	0	0
0	0	0	0	0	1
0	1	0	0	0	1
1	0	1	1	1	0
0	1	1	1	0	0
0	1	0	0	1	0

```

C:\Windows\System32\cmd.exe
max 3000.000000 ... 1.000000

[8 rows x 7 columns]
0 Kargo Gülnur Korkmaz İştirme Engelliyim--...
1 Kargo Teslimat Yapmamak İçin Elllerinden Ge...
2 Kargo Covid Önlemleri Alınmıyor--> PTT Kar...
3 Kargo Ürünü Adrese Getirmeme, Müşteri Hizm...
4 Kargo Kargom Yola Çıkıyor--> Şubeye gitt...
Name: icerik, dtype: object
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1887 entries, 0 to 1886
Data columns (total 8 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   id                                         1887 non-null   int64
1   icerik                                    1887 non-null   object
2   Gecikti veya Dağıtıma Çıkmadı            1887 non-null   int64
3   Evde yok notu düşüldü veya Kapiya Getirilmedi  1887 non-null   int64
4   Telefonlara Cevap Verilmedi             1887 non-null   int64
5   İade Süreci                              1887 non-null   int64
6   Teslim Alınmadı veya Teslim Edilmedi     1887 non-null   int64
7   Hijyen Kurallarına Uyulmadı             1887 non-null   int64
dtypes: int64(7), object(1)
memory usage: 110.6+ KB

HashingVectorizer yöntemi ile train edilen 1887 test verisinde yapılan toplam 11322 tah
minden 8691 kadarı doğrudur
doğruluk oranı= %76.76205617382088

d:\Users\TKPC\Desktop\multilabel kelime>

```

Şekil 4.13. Hashingvectorizer ile tahminlenen verilerin, anket bazlı etiket atamalarıyla arasındaki benzerlik oranı

Yapılan karşılaştırmalar sonucu 1887 eğitim verisi üzerinde yapılan tahminleme işlemi sonuçlarının, tahminlenen verilerdeki gerçek etiket atamaları ile arasında ortaya çıkan benzerlik oranı %76,76 olarak hesaplanmıştır. Tahminleme işlemi süresince 11322 tahmin yapılmış ve 8691 tanesi doğru tahmin edilmiştir. Bu ortalama ancak başarılı bir sonuçtur (Şekil 4.13.).

Elde edilen bu test sonuçlarının ardından 16333 adet olan toplam şikâyet içerikleri üzerinde bu modelin uygulamasına geçilmiştir. Uygulama sonuçlarına göre: “Gecikti veya Dağıtıma Çıkmadı” konusunda 7198, “Evde yok notu düşüldü veya Kapiya Getirilmedi” konusunda 5961, “Telefonlara Cevap Verilmedi” konusunda 6590, “İade Süreci” konusunda 4617, “Teslim Alınmadı veya Teslim Edilmedi” konusunda 7473 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış 5605 şikâyet tespit edilmiştir (Tablo 4.25.).

Tablo 4.25. HashingVectorizer ile anket bazlı tahminlenen toplam şikâyet verisinin kategorilere göre dağılımı

Kategori	Örnek Sayısı
Gecikti veya Dağıtıma Çıkmadı	7198
Evde yok notu düşüldü veya Kapiya Getirilmedi	5961
Telefonlara Cevap Verilmedi	6590
İade Süreci	4617
Teslim Alınmadı veya Teslim Edilmedi	7473
Hijyen Kurallarına Uyulmadı	5605

4.3. Kelime Bazlı Etiket Ataması ile Eğitilen Model ve Anket Yöntemi ile Eğitilen Model Sonuçlarının Değerlendirilmesi

Yapılan analizler sonucu elde edilen verilere göre “Gecikti veya Dağıtıma Çıkmadı” konusunda ortalama 7501, “Evde yok notu düşüldü veya Kapiya Getirilmedi” konusunda ortalama 6083, “Telefonlara Cevap Verilmedi” konusunda ortalama 6673, “İade Süreci” konusunda ortalama 3904, “Teslim Alınmadı veya Teslim Edilmedi” konusunda ortalama 6320 ve “Hijyen Kurallarına Uyulmadı” konusunda yapılmış ortalama 3660 şikâyet tespit edilmiştir (Tablo 4.26.).

Tablo 4.26. Toplam şikâyet verilerinin eğitilen yöntemlerine göre kategorilere dağılımı

Yöntem	Kategori	Gecikti veya Dağıtıma Çıkmadı	Evde yok notu düşüldü veya Kapiya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
	Bazlı						
Kelime TfidfVectorizer	Bazlı	8156	6176	6902	4189	5696	3104
Kelime CountVectorizer	Bazlı	8614	6336	7610	3431	5100	1078
Kelime HashingVectorizer	Bazlı	8075	5770	7001	3894	5375	4017
Anket TfidfVectorizer	Bazlı	7019	5991	6514	4751	7706	5196
Anket CountVectorizer	Bazlı	5941	6265	5418	2542	6568	2958
Anket HashingVectorizer	Bazlı	7198	5961	6590	4617	7473	5605
Ortalama		7501	6083	6673	3904	6320	3660

Bu durumda ortalama olarak en çok şikâyetin “Gecikti veya Dağıtıma Çıkmadı” konusunda yapıldığı ve en az şikâyetin ise “Hijyen Kurallarına Uyulmadı” konusunda yapıldığı çıkarımını yapmak mümkündür.

Tablo 4.27. Toplam şikâyet sayısının kargo firmaları bazında dağılımı

Firma	Yapılan Şikâyet Sayısı
A Firması	3261
B Firması	3892
C Firması	1793
D Firması	3817
E Firması	3569

Kargo firmalarına yapılan şikâyet sayısının kargo firmalarına göre dağılımı Tablo 4.27.'de gösterilmiştir. Kargo firmalarına yapılan şikâyetlerin kargo firmalarına ve alınan şikâyetlerin konusuna göre kategorilize edilerek dağılımı Tablo 4.28.'de gösterilmiştir.

Tablo 4.28. Toplam şikâyet verisi değerlerinin kargo firmaları bazında kategorilere göre dağılımı

	Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
A Firması	1549	1074	1337	961	1494	984
B Firması	1725	1413	1550	1127	1853	1183
C Firması	730	672	715	499	855	612
D Firması	1660	1383	1548	1102	1823	1218
E Firması	1354	1447	1364	1059	1678	1195

Kargo firmalarına yapılan şikâyetlerin her bir kargo firmasının aldığı şikâyet sayısına oranlanarak her bir kargo firmasının aldığı 100 şikâyete göre yüzde kaçının hangi kategorilere ait olduğu Tablo 4.29.'da gösterilmiştir.

Tablo 4.29. Toplam şikâyet verisi değerlerinin kargo firmaları bazında kategorilere göre yüzdelik dağılımı

	Gecikti veya Dağıtım Çıkmadı	Evde yok notu düşüldü veya Kapıya Getirilmedi	Telefonlara Cevap Verilmedi	İade Süreci	Teslim Alınmadı veya Teslim Edilmedi	Hijyen Kurallarına Uyulmadı
A Firması	47,50077	32,934683	40,99969	29,46949	45,81417	30,17479
B Firması	44,32169	36,305242	39,82528	28,95683	47,61048	30,39568
C Firması	40,71389	37,479085	39,8773	27,83045	47,68544	34,13274
D Firması	43,48965	36,232643	40,55541	28,87084	47,76002	31,90988
E Firması	37,9378	40,54357	38,21799	29,67218	47,01597	33,48277

Anket yöntemiyle etiket ataması ile eğitilen modele göre 1887 eğitim verisi üzerinde yapılan toplamda 33966 tahmine göre hazırlanan karışıklık matrisi Tablo 4.30.'da gösterilmiştir. Yöntemin doğruluk oranı %87 olarak hesaplanmıştır.

Tablo 4.30. Anket yöntemiyle etiket ataması karışıklık matrisi

Yapılan toplam tahmin sayısı = 33966	Anket yöntemiyle etiket ataması karışıklık matrisi		
	Tahmin: Yok	Tahmin: Var	Toplam
Gerçek: Yok	21631	805	22436
Gerçek: Var	3620	7910	11530
Toplam	25251	8715	
Doğruluk	0,869723		

BÖLÜM 5. TARTIŞMA VE SONUÇ

Sonuç olarak pandemi döneminde insanlar kargo firmalarından en çok kargolarının gecikmesinden veya dağıtıma çıkmamasından şikâyetçi olmuşlardır. En az şikâyet edildiği tespit edilen konu ise kargo çalışanlarının hijyen kurallarına uymaması olmuştur. Çalışmada kullanılan eğitim verisine etiketleme yapılması için, kelime bazlı etiket ataması ve anket yöntemi ile etiket ataması olmak üzere iki farklı yöntem kullanılmıştır. Etiketleme ve vektörleme yöntemlerine göre her bir kategori için tahminlenen şikâyet sayıları ve tüm yöntemlerle tahminlenen toplam şikâyet sayıları Tablo 5.1.'de gösterilmiştir. Bu tabloda gösterilen sonuçlara göre kelime bazlı etiket ataması yöntemiyle toplamda 100524 adet şikâyetin tahminlenebildiği ve anket yöntemiyle etiket ataması yöntemiyle de toplamda 104313 adet şikâyetin tahminlenebildiği görülmüştür. Çalışma içerisinde oluşturulan karışıklık matrislerine göre kelime bazlı etiket ataması yöntemiyle %92'lik bir doğruluk oranı elde edilirken, anket yöntemiyle etiket ataması yöntemiyle ise %87'lik bir doğruluk oranı elde edilmiştir. Bu sonuçlara bakarak; anket yöntemiyle etiket ataması ile daha fazla sayıda şikâyetin tahminlenebildiği ancak buna karşın doğruluk oranının azaldığı gözlemlenmiştir. Her iki etiketleme yönteminde de en yüksek doğruluk oranı %99 gibi oldukça yüksek oranlarla CountVectorizer vektörleme yaklaşımı olmuştur. Bunun sebebinin bu yöntemin çalışma prensibinden kaynaklandığı düşünülmektedir. CountVectorizer vektörleme yaklaşımı, belge içerisindeki kelimeleri sayarak vektörleme yapmaktadır. Eğitim verilerinin test edilmesi sürecinde eğitim ve test verisinin aynı örneklerden oluşmasından dolayı, yöntemin aynı örnekler üzerinde aynı kelime sayılarını bulmuş olması ve bu yüzden de yüksek isabetli tahminler yapmış olması mümkündür. Ancak bu sonucun yanıltıcı olabileceği ve bu yaklaşım ile oluşturulan model, eğitim verisinden farklı şikâyet verileri üzerinde çalıştırıldığında, aynı oranda yüksek doğrulukta tahminlerde bulunamayabileceği de göz önünde bulundurulmalıdır.

Tablo 5.1. Etiketleme ve vektörlendirme yöntemlerine göre her bir kategori için tahminlenen şikâyet sayıları

Etiketleme Yöntemi			
Şikâyet	Kelime Bazlı Etiket Ataması	Anket Yöntemi ile Etiket Ataması	Yöntem
Gecikti veya Dağıtıma	8156	7019	TfidfVectorizer
Çıkmadı	8614	5941	CountVectorizer
	8075	7198	HashingVectorizer
Evde yok notu düşüldü	6176	5991	TfidfVectorizer
veya Kapiya Getirilmedi	6336	6265	CountVectorizer
	5770	5961	HashingVectorizer
Telefonlara Cevap	6902	6514	TfidfVectorizer
Verilmedi	7610	5418	CountVectorizer
	7001	6590	HashingVectorizer
İade Süreci	4189	4751	TfidfVectorizer
	3431	2542	CountVectorizer
	3894	4617	HashingVectorizer
Teslim Alınmadı veya	5696	7706	TfidfVectorizer
Teslim Edilmedi	5100	6568	CountVectorizer
	5375	7473	HashingVectorizer
Hijyen Kurallarına	3104	5196	TfidfVectorizer
Uyulmadı	1078	2958	CountVectorizer
	4017	5605	HashingVectorizer
Toplam	100524	104313	Tüm Yöntemler

Etiketleme ve vektörlendirme yöntemlerine göre her bir kategori için tahminlenen şikâyetlerin doğruluk oranları Tablo 5.2.'de gösterilmiştir.

Tablo 5.2. Tahminlenen şikâyetlerin doğruluk oranları

Doğruluk Oranları	Anket Yöntemi			Kelime Yöntemi		
	TfidfVecto rizer	CountVect orize	HashingVect orizer	TfidfVecto rizer	CountVect orize	HashingVect orizer
Gecikti veya Dağıtıma Çıkmadı	0,8574	0,9979	0,7780	0,8590	0,9977	0,7766
Evde yok notu düşüldü veya Kapiya Getirilmedi	0,8739	0,9963	0,8113	0,9423	0,9997	0,9033
Telefonlara Cevap Verilmedi	0,8362	0,9921	0,7472	0,9333	1,0000	0,8620
İade Süreci	0,8256	0,9857	0,7843	0,9620	1,0000	0,9363
Teslim Alınmadı veya Teslim Edilmedi	0,8808	0,9984	0,7467	0,8886	0,9933	0,8516
Hijyen Kurallarına Uyulmadı	0,8151	0,9899	0,7382	0,8833	0,9997	0,7843

Kelime bazlı etiket ataması ve anket yöntemiyle etiket ataması ile eğitilen modellere göre eğitim verileri üzerinde yapılan tahminlere göre hazırlanan karışıklık matrisleri Tablo 5.3.'de gösterilmiştir.

Tablo 5.3. Tahminlere göre hazırlanan karışıklık matrisleri

Kelime Bazlı Etiketleme					Anket Yöntemiyle Etiketleme				
TfidfVectorizer					TfidfVectorizer				
Tahmin					Tahmin				
		Yok	Var	Toplam			Yok	Var	Toplam
Gerçek	Yok	11059	519	11578	Gerçek	Yok	6950	252	7202
	Var	1075	5341	6416		Var	1467	2653	4120
	Toplam	12134	5860	17994		Toplam	8417	2905	11322
	Doğruluk	0,911415				Doğruluk	0,848172		
HashingVectorizer					HashingVectorizer				
Tahmin					Tahmin				
		Yok	Var	Toplam			Yok	Var	Toplam
Gerçek	Yok	10356	879	11235	Gerçek	Yok	6339	553	6892
	Var	1778	4981	6759		Var	2078	2352	4430
	Toplam	12134	5860	17994		Toplam	8417	2905	11322
	Doğruluk	0,85234				Doğruluk	0,767621		
CountVectorizer					CountVectorizer				
Tahmin					Tahmin				
		Yok	Var	Toplam			Yok	Var	Toplam
Gerçek	Yok	12113	8	12121	Gerçek	Yok	8342	0	8342
	Var	21	5852	5873		Var	75	2905	2980
	Toplam	12134	5860	17994		Toplam	8417	2905	11322
	Doğruluk	0,998388				Doğruluk	0,993376		

Türkiye’deki önde gelen kargo firmalarının, elde edilen sonuçlara göre en başarılı ve en başarısız yönleri, konularına göre aldıkları şikâyet sayılarının, aldıkları toplam şikâyet sayılarına oranlanıp karşılaştırılarak her bir 100 şikâyet için yüzde bazında yorumlandığında; kargoların zamanında teslimi veya dağıtımına çıkması konusunda en başarılı kargo firması bu konuda en az şikâyet alan E Firması, en başarısız kargo firması ise bu konuda en çok şikâyet alan A Firması olmuştur. Evde yok notu düşülmesi veya kargoların kapıya getirilmemesi konusunda en başarılı kargo firması A Firması, en başarısız kargo firması ise E Firması olmuştur. Telefonlara cevap verme konusunda en başarılı kargo firması E Firması, en başarısız kargo firması ise A Firması olmuştur. Kargoların iade edilmesi süreci konusunda en başarılı kargo firması C Firması, en başarısız kargo firması ise E Firması olmuştur. Kargoların teslim alınmaması veya teslim edilmemesi konusunda en başarılı kargo firması A

Firması, en başarısız kargo firması ise D Firması olmuştur. Ve son olarak da kargoların hijyen kurallarına uyularak teslim edilmesi ve genel hijyen kurallarının dikkate alınması konusunda en başarılı kargo firması A Firması, en başarısız kargo firması ise C Firması olmuştur.

Bu çalışmanın sonuçları, Türkiye’deki önde gelen kargo firmalarının pandemi dönemini nasıl yönettiklerini ortaya koymaktadır. Kargo firmaları müşteri memnuniyetini artırmak ve daha az şikâyet almak istiyorlarsa, kendilerinin ve rakiplerinin durumlarını gözlemleyerek eksik yönlerini kapatmaya çalışmalı ve daha başarılı olmaya çabalamalıdır.

Çalışmadan elde edilen veriler incelendiğinde; şikâyet sayısının en yoğun olduğu dönemin, vaka sayısının ilk defa anlamlı ölçüde yükseliş gösterdiği dönemin hemen sonrası olduğu gözlemlenmiştir. Bunun sebebi olarak Türkiye’de yaşayan insanların birçoğunun pandemi dolayısıyla sağlık bakanlığınca yapılan “Evde Kal!” çağrılarını uyarak evde kalmayı tercih ettikleri, gıda, giyim, alışveriş gibi her türlü ihtiyaçlarını dışarıya çıkıp temin etmek yerine firmalardan telefon ya da internet aracılığıyla sipariş etme yoluna başvurdıkları ve sipariş edilen ürünlerin kargo firmalarınca getirilmesi sürecinde yaşanan olumsuzluklardan dolayı bu şikâyetlerin artış gösterdiği düşünülmektedir.

Çalışmada ulaşılan bir diğer bulgu da vaka sayılarının arttığı dönemlerde zaman ilerledikçe vaka yoğunluğunun artmasına karşın ters orantılı olarak şikâyet yoğunluğunun azalmasıdır. Bunun sebebi olarak da vatandaşların vaka yoğunluğunun arttığı dönemlerde, zaman ilerleyip pandemi şartlarına alıştıkça “Evde Kal!” çağrılarını giderek daha az uyması daha az online alışverişe yönelmesi ve bunun sonucu olarak daha az olumsuzluk yaşayarak daha az şikâyetle bulunması olarak düşünülmektedir. Aynı zamanda bu durumun bir diğer sebebi olarak, kargo firmalarının bu süreçte pandemi şartlarına uyum sağlamasıyla ve giderek daha iyi hizmet sunmasıyla şikâyet yoğunluğunu azaltmış olması da mümkündür.

Genel olarak değerlendirildiğinde, analiz bulgularına göre en çok şikâyetin kargoların dağıtımı sırasında yaşandığı ve gecikme yaşanan vaka sayısının yüksek olduğu gözlenmektedir. Özellikle pandemi dönemi içerisinde artış gösteren e-ticaret hacmine karşın kargo firmalarının aynı ölçüde gelişme gösteremeyerek yetersiz kalması, oluşan problemin asıl temel sebebi olabilir. Bu şekilde oluşan şikâyetlerin minimum seviyeye indirgenmesi için kargo firmalarının, artmakta olan iş hacmine göre şube sayısını çoğaltarak ve daha fazla çalışan ile bu sektördeki istihdamı artırarak cevap vermesi bir gerekliliktir.

Araştırma sonucuna göre elde edilen bir diğer bulgu ise şube çalışanlarının ve müşteri hizmetlerinin, saha personellerinde olduğu gibi artış gösteren iş hacmi karşısında kontrolsüzlük yaşadığı görülmüştür. Müşteri hizmetlerinin ilgisiz tutumu ve müşteriden gelen taleplerin dikkate alınmıyor olması (şubede teslim edilmesi ancak müşterilerin kapıda teslim istemesi gibi) dikkat çeken bir şikâyet faktörü olmuştur. İş yoğunluğundan kaynaklı olarak yetersiz eğitimler ile işe başlatılan personel ya da çalışanlar üzerinde oluşan tükenmişlik algısının temel olarak bu şikâyetlerin asıl kaynağı olduğu düşünülmektedir. Bu durumda, kargo firmalarının şirket içi verilen eğitimlere önem vermesi ve personeline yönelik motive edici unsurları dikkate alması, bu konulardan kaynaklanan şikâyetleri önemli ölçüde azaltabilir. Bu alanda yapılması planlanmakta olan çalışmalarda öğrenme verisinin daha fazla olmasının tahmin tutarlılığını olumlu yönde arttıracığı ve kelime ağırlıklarının hesaplanması sırasında çeşitli yöntemlerden faydalanılması veya diğer sınıflandırma algoritmalarının kullanılmasının da elde edilen sonuçları daha tutarlı hâle getireceği düşünülmektedir.

Literatürde yer alan, pandemi döneminde oluşturulmuş güncel çalışmalara göre E-ticaret sitelerinde var olan sipariş yoğunluğunun giderek arttığı ve buna bağlı olarak bu alandaki sorunların da büyük ölçüde arttığı belirtilmiştir. Bu çalışmada elde edilen sonuçlar da literatürdeki çalışmalarla benzer bulguları oluşturmuştur. Aynı zamanda yine literatüre göre pandemi döneminde ürünlerin kargoyla teslimatında yoğunluğa bağlı olarak büyük aksamalar yaşandığı bulgusu da yine bu çalışmanın sonuçlarıyla örtüşmektedir.

Araştırma kısıtları ve gelecek çalışmalar: Bu tez çalışmasında kullanılan verilerin yalnızca tek bir web sayfası üzerinden alınmış olması bu çalışmanın bir kısıtıdır. Daha farklı platformlardan elde edilecek verilerin kullanılması daha yeni sonuçlar doğurabilir. Aynı zamanda şikâyetlerde bulunan kişilerin demografik özellikleri ve aynı problemler için birden fazla şikâyet yazıp yazmadıkları veya yazılan şikâyetlerin doğru olup olmadığı tespit edilememiştir. Ayrıca anket yöntemi ile etiket ataması için eğitim verisinin modellenmesinde kullanılan ankette şikâyetleri kategorilerine ayıran katılımcıların, bu şikâyetleri gerçekten okuyup okumadıkları ve işaretledikleri kategorilerin doğru olup olmadığı bilinmemektedir. Modele eklenen yanlış kategorilendirmeler başarı sonucunu düşürmüş olabilir. Oluşan verinin artması ve teknolojinin zaman geçtikçe giderek gelişmesi bu alanda yapılabilecek yeni çalışma sahaları oluşturmaktadır. Çalışmada sınıflandırma algoritması olarak yalnızca Lojistik Regresyon yöntemi kullanılmıştır. Bunun sebebi diğer algoritmalar kullanıldığında elde edilen sonuçların tutarsız olmasıdır. Bu alanda yapılması planlanan çalışmaların öncelikle farklı platformları da araştırarak birden fazla veri kaynağı kullanmaları ve verilerin analizinde farklı yöntemleri de test etmeleri literatür için daha faydalı olacaktır.

Bu çalışmada kullanılan verinin kısıtlı olmasından dolayı elde edilen sonuçlar sınırlı olmuştur. Literatüre bakıldığında bu alanda yapılan çalışmalarda yöntem kısıtı olduğu görülmüştür. Bu tez çalışmasında kargo şikâyetlerinin değerlendirilmesinde, eğitim verisinin doğru bir şekilde kategorilendirilebilmesi için, kelime bazlı etiket ataması ve anket yöntemi ile etiket ataması olmak üzere iki farklı yöntem önerilmiştir. Çok etiketli sınıflandırma algoritmalarının kullanılacağı gelecek çalışmalarda, bu çalışmada önerilen yöntemlerin kullanılabilir olmasıyla, literatürde bu alanda bulunan boşluğun doldurulmasına katkı sağlanması amaçlanmıştır.

KAYNAKLAR

- Aksu, M. Ç. ve Karaman, E., (2020). FastText ve Kelime Çantası Kelime Temsil Yöntemlerinin Turistik Mekanlar İçin Yapılan Türkçe İncelemeler Kullanılarak Karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi* Sayı 20, S. 311-320,
- Aktaş, N. B., Aktaş, B., ve Akbıyık, A. (2020). Koronavirüs ‘ün (Covid-19) Türkiye ‘de E-Ticaret Mü ş terilerinin Kalite Alg ı s ı na Etkisi : Trendyol Örne ğ i. January 2021.
- Alabay, M. N. (2012). Müşteri Şikâyetleri Yönetimi. *International Journal of Management Economics and Business*, 8(16), 137–158.
- Alfonso, V., Boar, C., Frost, J., Gambacorta, L., ve Liu, J. (2020). Online appendix - E-commerce in the pandemic and beyond. 1–5.
- Altın, Z. (2020). Elderly People in Covid-19 Outbreak. *The journal of Tepecik Education and Research Hospital*, 30, 49–57. <https://doi.org/10.5222/terh.2020.93723>
- Analyticsvidhya.com. (2021). Commonly used Machine Learning Algorithms (with Python and R Codes). <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>, Erişim Tarihi: 15.05.2021
- Bell, S. J., ve Luddington, J. A. (2006). Coping with customer complaints. *Journal of Service Research*, 8(3), 221–233. <https://doi.org/10.1177/1094670505283785>
- Budak, F., ve Korkmaz, Ş. (2020). Covid-19 Pandemi Sürecine Yönelik Genel Bir Değerlendirme: Türkiye Örneği. *Sosyal Araştırmalar ve Yönetim Dergisi*, 1, 62–79. <https://doi.org/10.35375/sayod.738657>
- Cho, Y., Im, I., Hiltz, R., ve Fjermestad, J. (2002). An analysis of online customer complaints: Implications for Web complaint management. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2002-Janua(c), 2308–2317. <https://doi.org/10.1109/HICSS.2002.994162>
- Covid19.saglik.gov.tr. (2020). COVID-19 Bilgilendirme Platformu. <https://covid19.saglik.gov.tr/TR-66935/genel-koronavirus-tablosu.html>, Erişim Tarihi: 21.12.2020
- Covid19.saglik.gov.tr. (2022). COVID-19 Bilgilendirme Platformu. <https://covid19.saglik.gov.tr/TR-66935/genel-koronavirus-tablosu.html>, Erişim Tarihi: 15.05.2022
- Çelik, Ö. Ve Koç, B. C. (2021) TF-IDF, Word2vec ve Fasttext Vektör Model Yöntemleri ile Türkçe Haber Metinlerinin Sınıflandırılması. *DEUFMD*, 23(67), 121-127

- Dataschool.io. (2022). Simple guide to confusion matrix terminology. <https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/>., Erişim Tarihi: 11.02.2022
- Davidow, M. (2003). Organizational Responses to Customer Complaints: What Works and What Doesn't. *Journal of Service Research*, 5(3), 225–250. <https://doi.org/10.1177/1094670502238917>
- De Carvalho, A. C. P. L. F., ve Freitas, A. A. (2009). A tutorial on multi-label classification techniques. *Studies in Computational Intelligence*, 205, 177–195. https://doi.org/10.1007/978-3-642-01536-6_8
- Deepakumara, J., Heys, H. M., Venkatesan, R. (2001). FPGA implementation of MD5 hash algorithm. <https://doi.org/10.1109/CCECE.2001.933564>
- Demirdöğmez, M., Gültekin, N., ve Taş, H. Y. (2018). Türkiye’de E- Ticaret Sektörünün Yıllara Göre Gelişimi Development of E-Commerce Sector In Turkey By Years. <https://doi.org/10.26466/opus.453147>
- Demirdöğmez, M., Taş, H. Y., ve Gültekin, N. (2020). Koronavirüs’ün (Covid-19) E-Ticarete Etkileri. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 16(29), 1–1. <https://doi.org/10.26466/opus.734477>
- Deniz, A., ve Gödekmerdan, L. (2011). Müşterilerin Kargo Firmalarının Sunduğu Hizmetlere Yönelik Tutum ve Düşünceleri Üzerine Bir Araştırma. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 15(2), 379–396.
- Erdem, A. T., ve Akolaş, D. A. (2020). MÜŞTERİ MEMNUNİYETİNİN LOJİSTİK YÖNETİMİ BAĞLAMINDA İNCELENMESİ: BİR KARGO ŞİRKETİ ÖRNEĞİ Investigation of Customer Satisfaction in the Context of Logistics Management : The Case of A Cargo Company. 2(2), 139–159.
- Eticaret.gov.tr. (2022). 2020 ve 2021 Yılı E-Ticaret Verileri. <https://eticaret.gov.tr/istatistikler/>., Erişim Tarihi: 17.04.2022
- Ey.com. (2020). EY-Parthenon & ETİD COVID-19 Yönetici ve KOBİ Anketleri. https://www.ey.com/tr_tr/ey-turkiye-yayinlar-raporlar/ey-parthenon-etid-covid-19-yoneticisi-ve-kobi-anketleri/., Erişim Tarihi: 19.11.2020
- Gadde, S., Lakshmanarao, A., Satyanarayana, S. (2021). SMS Spam Detection using Machine Learning and Deep Learning Techniques. <https://doi.org/10.1109/ICACCS51430.2021.9441783>
- Goyal, R. (2021). Evaluation of rule-based, CountVectorizer, and Word2Vec machine learning models for tweet analysis to improve disaster relief. <https://doi.org/10.1109/GHTC53159.2021.9612486>
- Jimenez, G., Novoa, L., Ramos, L., Martinez, J., ve Alvarino, C. (2018). Diagnosis of Initial Conditions for the Implementation of the Integrated Management System in the Companies of the Land Cargo Transportation in the City of Barranquilla (Colombia). Springer International Publishing. <https://doi.org/10.1007/978-3-319-92285-0>
- Kwartler, T. (2017). What is Text Mining? Text Mining in Practice with R, 1–15. <https://doi.org/10.1002/9781119282105.ch1>

- Lisińska-Kuśnierz, M., ve Gajewska, T. (2014). Customer satisfaction with the quality of the logistic services. *LogForum* 10 (1), 13-19.
- Machinelearningmastery.com. (2022). One-vs-Rest and One-vs-One for Multi-Class Classification. [https://machinelearningmastery.com/one-vs-rest-and-one-vs-one-for-multi-class-classification/#:~:text=One%2Dvs%2Drest%20\(OvR%20for%20short%2C%20also%20referred,into%20multiple%20binary%20classification%20problems.,](https://machinelearningmastery.com/one-vs-rest-and-one-vs-one-for-multi-class-classification/#:~:text=One%2Dvs%2Drest%20(OvR%20for%20short%2C%20also%20referred,into%20multiple%20binary%20classification%20problems.,) Erişim Tarihi: 23.03.2022
- Naqa, I. El, ve Murphy, M. J. (2015). Machine Learning in Radiation Oncology. *Machine Learning in Radiation Oncology*, 3–11. <https://doi.org/10.1007/978-3-319-18305-3>
- Neptune.ai. (2022). Vectorization Techniques in NLP [Guide]. <https://neptune.ai/blog/vectorization-techniques-in-nlp-guide.>, Erişim Tarihi: 16.04.2022
- Nikolaeva, I. G., Rudakova, E. N., Dmitrieva, O. A., ve Vlasov, A. V. (2020). Digitalization of the EAEU Transport and Logistics Sector and Its Role in Improving the Euro-Asian Cargo Transportation. *416(Iccese)*, 1296–1302.
- Onurlubaş, E., ve Gümüş, N. (2020). Kargo Firmalarının Hizmet Kalitesinin Müşteri Memnuniyeti Üzerine Etkisinin Servqual Ölçeği ile İncelenmesi.
- Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Prespectives and Prospects, COMITCon 2019*, 35–39. <https://doi.org/10.1109/COMITCon.2019.8862451>
- Resmigazete.gov.tr. (2022). COVID-19 Kapsamında Kamu Çalışanlarına Yönelik Tedbirler. <https://www.resmigazete.gov.tr/eskiler/2022/04/20220409-10.pdf.>, Erişim Tarihi: 11.05.2022
- Sabuhari, R., ve Irawanto, D. W. (2020). The Importance of Flexibility of Human Resources and Employee Performance on Logistic Industry in the Eastern part of Indonesian Region : A Literature Review. *144(Afbe 2019)*, 205–210.
- Schafer, J. Ben, Konstan, J., ve Riedl, J. (1999). Recommender Systems in E-Commerce. *EC '99: Proceedings of the 1st ACM conference on Electronic commerce* November 1999, 158–166. <https://doi.org/doi.org/10.1145/336992.337035>
- Scrapy.org. (2021). Scrapy 2.6 documentation page getting help. [https://docs.scrapy.org/en/latest/.](https://docs.scrapy.org/en/latest/), Erişim Tarihi: 18.04.2021
- Sikayetvar.com. (2022). Şikayetvar Hakkımızda. <https://www.sikayetvar.com/hakkimizda.>, Erişim Tarihi: 13.05.2022
- Songur, G., ve Büyükkeklik, A. (2016). Kargo Hizmet Sağlayıcılarında Hizmet Kalitesi Ve Kurumsal Müşteri Memnuniyeti: Konya İli Örneği. *Uluslararası İktisadi ve İdari Bilimler Dergisi*, 2(1), 103–119.
- Szyma, P., ve Kajdanowicz, T. (2016). scikit-multilearn: A scikit-based Python environment for performing multi-label classification. *1*, 1–15.

- Tan, A.-H. (1999). Text Mining: The state of the art and the challenges. Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases, 8, 65–70. <https://doi.org/10.1.1.38.7672>
- Towardsdatascience.com. (2022). Perfect Recipe for Classification Using Logistic Regression. <https://towardsdatascience.com/the-perfect-recipe-for-classification-using-logistic-regression-f8648e267592#:~:text=Logistic%20Regression%20is%20a%20statistical,dependent%20and%20the%20independent%20variable.>, Erişim Tarihi: 19.02.2022
- Tran, L. T. T. (2021). Managing the effectiveness of e-commerce platforms in a pandemic. *Journal of Retailing and Consumer Services*, 58(September 2020), 102287. <https://doi.org/10.1016/j.jretconser.2020.102287>
- Wilks, Y. (1996). Natural Language Processing. *Communications of the ACM*, 39(1), 60–62. <https://doi.org/10.1145/234173.234180>
- Worldometers.info. (2022). COVID-19 CORONAVIRUS PANDEMIC. <https://www.worldometers.info/coronavirus/>, Erişim Tarihi: 11.04.2022
- Yangping, Z. (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China.
- Zhang, W., Yoshida, T., ve Tang, X. (2011). A comparative study of TF*IDF, LSI and multi-words for text classification. *Expert Systems with Applications*, 38(3), 2758–2765. <https://doi.org/10.1016/j.eswa.2010.08.066>
- ZU, Z. Y., DI Jiang, M., XU, P. P., CHEN, W., NI, Q. Q., LU, G. M., ve ZHANG, L. J. (2020). Coronavirus Disease 2019 (COVID-19): A Perspective from China. *Radiology*, 296(2), E15–E25. <https://doi.org/10.1148/radiol.2020200490>

EKLER

EK A. Kodlar

```
import scrapy
from scrapy import Selector
import sqlite3
import csv
import os
import time

class MySpider(scrapy.Spider):
    name = "link"
    marka="kargomarkasi"
    kelime=""
    sonsayfa=20

    start_urls = []
    for i in range (1,sonsayfa+1):
        start_urls = start_urls + ["https://www.sikayetvar.com/"+str(marka)+"/"+kelime+"?page="+str(i)]
    def parse(self, response):
        with open("linkler.txt", "a", encoding="utf-8") as file:
            time.sleep(0.2)
            liste = response.xpath(
                '/html/body/div[1]/main/div/div/section[2]/div[1]/div[2]/article/section/h2/a/@href').extract()
            liste = str(liste).replace("'", " ", ",")
            liste = str(liste).replace("[", " ")
            liste = str(liste).replace("]", " ")
            liste = str(liste).replace(" ", "https://www.sikayetvar.com")

            file.write(str(liste))
        return 0
```

```
import scrapy
from scrapy import Selector
import sqlite3
import csv
import os

exists = os.path.isfile('linkler.txt')
if exists:
    linkfile = open("linkler.txt", "r")
    alinanLinkler = linkfile.read()
    alinanLinkler = alinanLinkler[1:len(alinanLinkler) - 1]
    with open("linkler.txt", "w", encoding="utf-8") as file:
        file.write(str(alinanLinkler))
    file.close()

    crimefile = open("linkler.txt", 'r')
    reader = csv.reader(crimefile)
    alinanLinkler = [row for row in reader]

else:
    linkDizi=[]
```

```

print(linkDizi)
say=0

conn = sqlite3.connect('database.db')
c = conn.cursor()
c.execute("CREATE TABLE IF NOT EXISTS "+sikayetvar+"(id TEXT, marka TEXT, kisi TEXT, tarih
TEXT, link TEXT, okunma TEXT, durum TEXT, baslik TEXT, icerik TEXT)")

class MySpider(scrapy.Spider):
    name = "icerik"
    start_urls = linkDizi

    def parse(self, response):
        global say
        import time
        time.sleep(0.4)
        id=str(response.xpath("//*[@id='complaint-id']/@data-id").extract_first())
        marka = str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[1]/div[1]/span/a/text()').extract_first())
        kisi= str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[2]/article/div[2]/div[1]/div/span/a/text()').extract_first())
        tarih= str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[2]/article/div[2]/div[1]/div/span/span/@title').extract_fir
        st())
        link= str(response.xpath('/html/head/link[9]/@href').extract_first())
        okunma= str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[2]/article/div[1]/span/span/text()').extract_first())
        durum= str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[1]/div[1]/ul/li[2]/span[2]/@title').extract_first())
        baslik = str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[2]/article/div[1]/h1/text()').extract_first())
        icerik= str(response.xpath(
        '/html/body/div[1]/main/div/div/section[2]/section[2]/article/div[2]/div[2]/div//text()').extract()).replace(",
        ", "").replace(" ", "").replace(";", "").replace(":", "").replace("?", "").replace("!", "").replace("&", "").replace("<", "").replace(">", "").replace(""", "").replace("'", "").replace("\\n", "\n").replace("\\xa0", "")

        say = say + 1
        print(str(say)+" "+baslik+"\n\n")
        with open("aciklamalar.txt", "a", encoding="utf-8") as file2:
            file2.write(str(say)+" "+id+"\n ID: "+id+"\n MARKA: "+marka+"\n KİŞİ: "+kisi+"\n
            TARİH: "+tarih+"\n LİNK: "+link+"\n GÖRÜNTÜLENME: "+okunma+"\n DURUM:
            "+durum+"\n BAŞLIK: "+baslik+"\n İÇERİK: "+icerik+
            "\n\n")

        c.execute('SELECT * FROM '+sikayetvar+' ')
        c.execute("INSERT INTO "+sikayetvar+" VALUES (?,?,?,?,?,?,?)",(str(id), str(marka),
        str(kisi), str(tarih), str(link), str(okunma), str(durum), str(baslik), str(icerik)))
        conn.commit()

```

```

import xlwt
from openpyxl import Workbook, load_workbook
from xlwt import Workbook

wb = Workbook()
sheet1 = wb.add_sheet('Sheet 1')

wb2 = load_workbook("Tüm kargo verileri.xlsx")
ws = wb2.active

sheet1.write(0, 0, "İçerik")
toplamkelime=0
toplamicerik=""
#satirsayisi= 16335
satirsayisi= 1889

noktalama=" , ; . \ " ' : - ! % ( ) [ ] * / ? ‘ ’ “ ” + _ = % # £ ^ $ ½ & { } \ < > | ® ¢ "

numbers="0 1 2 3 4 5 6 7 8 9"

stopwords="""ve,veya,yada,ile,lakin,ama,veyahut,zaten,oysa,ki,oysaki,hep,şimdi,diyor,tarih,derdi,öyleyse,
öyle,yine,en,az,ay,
,yeniden,yu,sinde,işlem,sürekli,ayrıca,zaten,orum,çok,çoktan,sı,de,da,ta,te,tam,nasıl,itibaren,sız,siz,ben,biz
,ıyorlar,
,ıyor,edi,nuz,mu,diyor,önce,dedi,geldi,tamam,merhaba,bir,üzerinden,yaptım,edip,lerinin,defa,acaba,tarih,b
aşlık,bu,@,®,
,rağmen,mı,mı,mu,mü,km,ya,şu,bu,o,ben,sen,biz,siz,onlar,dan,korona,corona,covid,pandemi,virüs,an,hem,
un,in,ın,on,ön,ün,
,a,b,c,ç,d,e,f,g,ğ,i,j,k,l,m,n,o,ö,p,r,s,ş,t,u,ü,v,y,z,
"""

kelimekokleri="""korona,corona,covid,pandemi,virüs,dönem,şube,ücret,yalan,oyalayama,numara,hata,uya
rı,temsilci,not,önlem,
,dağıtım,trendyol,hepsiburada,n l l,gecikme,telefon,iade,parça,sipariş,huysuz,laubali,kayıp,hasar,adres,dem
et,kargo,kurye,yetkili,
,şube,işletme,teslimat,teslim,çalışan,personel,şirket,memnuniyet,eleman,sıramatik,ortam,bağırma,saygısız,
değiştirmede,
,şikayet,alışverişi,ürün,tarih,ankara,istanbul,bursa,paket,engel,kapı,apartman,süreç,gönderici,irsaliye,fatura,
iş,koruma,
,sistem,aktarma,cevap,internet,ödeme,hizmet,müşteri,para,sorumsuz,imza,meşgul,mağdur,kolay,kelime,ekr
an,süreç,sürec,
,maskeli,maskesiz,eldivenli,eldivensiz,zil,mesafe,ev,dedi,yok,mesaj,tavır,arkadaş,merhaba,

,edilmek,edild,edilec,ediler,edilen,edili,edilmi,edilmes,
,getirmek,getiri,getird,getirmes,getire,
,getirmemek,getirmed,getirmey,getirmemes,
,çıkılmamak,çıkılmamas,çıkılmıy,çıkımad,çıkımay,çıkımamam,çıkımamı,
,çıkımak,çıkılmış,çıkıac,çıkıab,çıkıy,çıkıktı,çıkımas,
,edilmemek,edilmed,edilmiy,edilem,edilmey,
,etmek,ediy,ederi,edere,ederl,ederi,edece,edebil,etmel,edipt,edipd,etmeli,edişi,edişl,edece,etti,etmiş,
,etmemek,etmemi,etmemel,etmemes,etmey,etmiy,etmed,edem,
,bırakmak,bırakt,bırakı,bıraka,bıraks,bırakc,bırakç,bırakmal,bırakmas,bırakmad,bırakmış,
,göndermek,gönderiy,gönderil,gönderiş,göndermiş,gönderd,göndermel,göndere,

```

,göndermemek,göndermiy,göndermed,göndermey,göndermemel,göndereme,
 ,gitmek,gidebilm,gidebilen,gidebild,gidec,gidiy,gitt,gidiş,gitmiş,gitmes,giden,
 ,gitmemek,gitmemes,gidebileme,gideme,gitmiy,gitmey,gitmed,gidemey,
 ,gelmek,gelmemes,gelebileme,geleme,gelmiy,gelmey,gelmed,gelemey,
 ,gelmek,gelebilm,gelebilen,gelebild,gelec,geliy,geld,geliş,gelmiş,gelmes,gelen,
 ,getirme,getird,getiri,getire,getirc,getirs,getirme,
 ,getirmeme,getirmed,getirmey,getirmi,getirmic,getirmes,getirmeme,
 ,çalışma,çalışar,çalışı,çalışab,çalışam,çalışt,çalışmı,çalışma
 ,bekletmek,bekleti,bekletmel,bekletmes,bekletmey,bekletmed,bekletmi,beklett,beklete,
 ,beklemek,bekli,bekley,bekliy,beklemiş,bekled,bekler,beklemel,beklemes,beklend,
 ,beklememek,beklemed,beklemey,beklemiy,beklemed,beklemez,beklememel,beklememi,beklenm,
 ,vermek,verd,vermiş,verip,verer,veren,verec,versi,verse,veriy,verebil,verir,verilmiş,
 ,vermemek,vermed,vermey,vermemi,vermes,vermemes,vermiy,veremed,verilmeme,verilmemi,verilmed,ve
 rilmey,verilmes,verilmem,verilmes,
 ,geçmek,geçeb,geçti,geçmel,geçen,geçer,geçip,geçiş,geçmiş,geçil,geçirmiş,geçiri,geçire,geçmes,geçmel,
 ,geçmemek,geçeme,geçem,geçmiy,geçmey,geçmed,geçmemi,geçmiyor,geçirmem,geçirmiy,geçirmey,geçir
 med,geçirmeme,geçmemel,geçmemes,
 ,aramak,aramal,aradı,arıy,aray,aramı,
 ,açmamak,açmıy,açmad,açmay,açılmad,açılmıy,
 ,parmak,parmağ,kağıd,kağıt,
 ,değişiklik,değişiklik,kayb,kayıp,
 ,ulaşılama,ulaşılama,ulaşıl,ulaşıp,ulaşmad,ulaşmam,ulaşama,ulaştı,ulaşab,ulaştırılan,ulaştırılam,ulaşamı,
 ,yapmamak,yapmay,yapmıy,yapmad,yapmak,yapt,yapıy,yapıl,yapılmamak,yapılmamı,yapılmad,yapılmı,y
 apılmak,yapıld,yapılı,
 ,olmak,oldu,olac,olab,olan,olmuş,oluy,olur,ols,
 ,olmamak,olmad,olmay,olmuy,olama,olmamas,
 ,yapıp,yapamıy,yapan,yapılab,yapılan,yapılmış,yapılmay,yapabil,yapmış,
 ,duymuy,duymad,duymama,duymay,duyama,duys,
 ,görmüy,görmed,görmeme,görmey,göremed,görs,
 ,çalma,çalmak,çald,çalıy,çalara,çalıp,çals,çalmama,çalmamak,çalmad,çalmay,çalmas,çalanl,çalmı,

""

""

```
icerik = icerik.replace(str("."), " dot ")
icerik = icerik.replace(str(","), " comma ")
icerik = icerik.replace(str("!"), " exclamation ")
icerik = icerik.replace(str("?"), " questionmark ")
```

for number in range (0,11):

```
    icerik = icerik.replace(" "+str(number)+" ", " number ")
```

""

```

for n in range(2, satirsayisi):
    icerik = ws["B" + str(n) + ""].value + "--> " + ws["C" + str(n) + ""].value

    icerik = icerik + " "
    icerik = icerik.replace("İ", "i")
    icerik = icerik.replace("I", "ı")
    icerik = icerik.replace("Ö", "ö")
    icerik = icerik.replace("Ü", "ü")
    icerik = icerik.replace("Ç", "ç")
    icerik = icerik.replace("Ş", "ş")
    icerik = icerik.lower()

for nokta in noktalama.split(" "):
    icerik = icerik.replace(nokta, " ")

for number in numbers.split(" "):
    icerik = icerik.replace(number, " ")

sadeicerik=""
kelimeler=icerik.split()
for kelime in kelimeler:

    for kok in kelimekokleri.split(","):
        if kelime.startswith(kok):
            kelime = kelime.replace(kelime, kok)
    """"

    for kok in kelimekokleri.split(","):
        if kok in kelime:
            kelime = kelime.replace(kelime, kok)
    """"

    for baglac in stopwords.split(","):
        if baglac == kelime:
            kelime = kelime.replace(kelime, " ")

sadeicerik = sadeicerik + kelime + " "

sadeicerik = sadeicerik.replace("değişiklik ", "değişiklik ")
sadeicerik = sadeicerik.replace("değişikliğ ", "değişiklik ")
sadeicerik = sadeicerik.replace("kağıd ", "kağıt ")
sadeicerik = sadeicerik.replace("parmağ ", "parmak ")

```

```

sadeicerik = sadeicerik.replace("duymuy ", "duymamak ")
sadeicerik = sadeicerik.replace("duymad ", "duymamak ")
sadeicerik = sadeicerik.replace("duymama ", "duymamak ")
sadeicerik = sadeicerik.replace("duymay ", "duymamak ")
sadeicerik = sadeicerik.replace("duyama ", "duymamak ")
sadeicerik = sadeicerik.replace("duys ", "duymamak ")

sadeicerik = sadeicerik.replace("görmüy ", "görmemek ")
sadeicerik = sadeicerik.replace("görmed ", "görmemek ")
sadeicerik = sadeicerik.replace("görmeme ", "görmemek ")
sadeicerik = sadeicerik.replace("görmey ", "görmemek ")
sadeicerik = sadeicerik.replace("göremed ", "görmemek ")
sadeicerik = sadeicerik.replace("görs ", "görmemek ")

sadeicerik = sadeicerik.replace("açmıy ", "açmamak ")
sadeicerik = sadeicerik.replace("sürec ", "süreç ")
sadeicerik = sadeicerik.replace("kayb ", "kaybolmak ")
sadeicerik = sadeicerik.replace("kayıp ", "kaybolmak ")
sadeicerik = sadeicerik.replace("değiştirmed ", "değiştirmemek ")
sadeicerik = sadeicerik.replace("koruma ", "korumak ")

sadeicerik = sadeicerik.replace("bekleti ", "bekletmek ")
sadeicerik = sadeicerik.replace("bekletmel ", "bekletmek ")
sadeicerik = sadeicerik.replace("bekletmes ", "bekletmek ")
sadeicerik = sadeicerik.replace("bekletmey ", "bekletmek ")
sadeicerik = sadeicerik.replace("bekletmed ", "bekletmek ")
sadeicerik = sadeicerik.replace("bekletmi ", "bekletmek ")
sadeicerik = sadeicerik.replace("beklett ", "bekletmek ")
sadeicerik = sadeicerik.replace("beklete ", "bekletmek ")
sadeicerik = sadeicerik.replace("bekletmek ", "bekletmek ")

sadeicerik = sadeicerik.replace("beklend ", "beklemek ")
sadeicerik = sadeicerik.replace("beklemes ", "beklemek ")
sadeicerik = sadeicerik.replace("beklemel ", "beklemek ")
sadeicerik = sadeicerik.replace("bekler ", "beklemek ")
sadeicerik = sadeicerik.replace("bekled ", "beklemek ")
sadeicerik = sadeicerik.replace("beklemiş ", "beklemek ")
sadeicerik = sadeicerik.replace("bekliy ", "beklemek ")
sadeicerik = sadeicerik.replace("bekley ", "beklemek ")
sadeicerik = sadeicerik.replace("bekli ", "beklemek ")
sadeicerik = sadeicerik.replace("beklemek ", "beklemek ")

sadeicerik = sadeicerik.replace("beklenm ", "beklememek ")
sadeicerik = sadeicerik.replace("beklememi ", "beklememek ")
sadeicerik = sadeicerik.replace("beklememel ", "beklememek ")
sadeicerik = sadeicerik.replace("beklemez ", "beklememek ")
sadeicerik = sadeicerik.replace("beklemed ", "beklememek ")
sadeicerik = sadeicerik.replace("beklemiy ", "beklememek ")
sadeicerik = sadeicerik.replace("beklemey ", "beklememek ")
sadeicerik = sadeicerik.replace("beklemed ", "beklememek ")
sadeicerik = sadeicerik.replace("beklememek ", "beklememek ")

```

```

sadeicerik = sadeicerik.replace("çalmama ", "çalma ")
sadeicerik = sadeicerik.replace("çalmas ", "çalma ")
sadeicerik = sadeicerik.replace("çalmamak ", "çalma ")
sadeicerik = sadeicerik.replace("çalmad ", "çalma ")
sadeicerik = sadeicerik.replace("çalmıy ", "çalma ")
sadeicerik = sadeicerik.replace("çalmay ", "çalma ")
sadeicerik = sadeicerik.replace("çalma ", "çalma ")
sadeicerik = sadeicerik.replace("çalmak ", "çalma ")
sadeicerik = sadeicerik.replace("çald ", "çalma ")
sadeicerik = sadeicerik.replace("çalıy ", "çalma ")
sadeicerik = sadeicerik.replace("çalara ", "çalma ")
sadeicerik = sadeicerik.replace("çalıp ", "çalma ")
sadeicerik = sadeicerik.replace("çals ", "çalma ")
sadeicerik = sadeicerik.replace("çalanl ", "çalma ")
sadeicerik = sadeicerik.replace("çalımı ", "çalma ")
sadeicerik = sadeicerik.replace("çalma ", "çalmak ")

sadeicerik = sadeicerik.replace("çalışma ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışar ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışı ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışab ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışam ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışt ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışımı ", "çalışma ")
sadeicerik = sadeicerik.replace("çalışma ", "çalışmak ")

sadeicerik = sadeicerik.replace("getirme ", "getirmeme ")
sadeicerik = sadeicerik.replace("getirmed ", "getirmeme ")
sadeicerik = sadeicerik.replace("getirmey ", "getirmeme ")
sadeicerik = sadeicerik.replace("getirmi ", "getirmeme ")
sadeicerik = sadeicerik.replace("getirmic ", "getirmeme ")
sadeicerik = sadeicerik.replace("getirmes ", "getirmeme ")
sadeicerik = sadeicerik.replace("getirmeme ", "getirmemek ")

sadeicerik = sadeicerik.replace("getirme ", "getirme ")
sadeicerik = sadeicerik.replace("getird ", "getirme ")
sadeicerik = sadeicerik.replace("getiri ", "getirme ")
sadeicerik = sadeicerik.replace("getire ", "getirme ")
sadeicerik = sadeicerik.replace("getirc ", "getirme ")
sadeicerik = sadeicerik.replace("getirs ", "getirme ")
sadeicerik = sadeicerik.replace("getirme ", "getirmek ")

sadeicerik = sadeicerik.replace("yapmamak ", "yapmamak ")
sadeicerik = sadeicerik.replace("yapmay ", "yapmamak ")
sadeicerik = sadeicerik.replace("yapmıy ", "yapmamak ")
sadeicerik = sadeicerik.replace("yapmad ", "yapmamak ")
sadeicerik = sadeicerik.replace("yapmak ", "yapmak ")
sadeicerik = sadeicerik.replace("yapt ", "yapmak ")
sadeicerik = sadeicerik.replace("yapıy ", "yapmak ")
sadeicerik = sadeicerik.replace("yapıl ", "yapmak ")
sadeicerik = sadeicerik.replace("yapıp ", "yapmak ")
sadeicerik = sadeicerik.replace("yapılmamak ", "yapılmamak ")
sadeicerik = sadeicerik.replace("yapılmamı ", "yapılmamak ")
sadeicerik = sadeicerik.replace("yapılmad ", "yapılmamak ")

```

```

sadeicerik = sadeicerik.replace("yapılmı ", "yapılmamak ")
sadeicerik = sadeicerik.replace("yapılmay ", "yapılmamak ")
sadeicerik = sadeicerik.replace("yapılmak ", "yapılmak ")
sadeicerik = sadeicerik.replace("yapıld ", "yapılmak ")
sadeicerik = sadeicerik.replace("yapılı ", "yapılmak ")
sadeicerik = sadeicerik.replace("yapamıy ", "yapamamak ")
sadeicerik = sadeicerik.replace("yapan ", "yapabilmek ")
sadeicerik = sadeicerik.replace("yapılabilir ", "yapabilmek ")
sadeicerik = sadeicerik.replace("yapılan ", "yapabilmek ")
sadeicerik = sadeicerik.replace("yapılmış ", "yapabilmek ")
sadeicerik = sadeicerik.replace("yapabil ", "yapabilmek ")
sadeicerik = sadeicerik.replace("yapmış ", "yapmak ")

sadeicerik = sadeicerik.replace("ulaşılama ", "ulaşılamamak ")
sadeicerik = sadeicerik.replace("ulaşılama ", "ulaşılamamak ")
sadeicerik = sadeicerik.replace("ulaşıl ", "ulaşılmak ")
sadeicerik = sadeicerik.replace("ulaşıp ", "ulaşmak ")
sadeicerik = sadeicerik.replace("ulaşmad ", "ulaşılmamak ")
sadeicerik = sadeicerik.replace("ulaşmam ", "ulaşılmamak ")
sadeicerik = sadeicerik.replace("ulaşama ", "ulaşılamamak ")
sadeicerik = sadeicerik.replace("ulaşamı ", "ulaşılamamak ")
sadeicerik = sadeicerik.replace("ulaştı ", "ulaşmak ")
sadeicerik = sadeicerik.replace("ulaşab ", "ulaşmak ")
sadeicerik = sadeicerik.replace("ulaştırılan ", "ulaşmak ")
sadeicerik = sadeicerik.replace("ulaştırılma ", "ulaşılamamak ")
sadeicerik = sadeicerik.replace("ulaşılmak ", "ulaşmak ")
sadeicerik = sadeicerik.replace("ulaşılamamak ", "ulaşmamak ")

sadeicerik = sadeicerik.replace("getirmek ", "getirmek ")
sadeicerik = sadeicerik.replace("getiri ", "getirmek ")
sadeicerik = sadeicerik.replace("getird ", "getirmek ")
sadeicerik = sadeicerik.replace("getirmes ", "getirmek ")
sadeicerik = sadeicerik.replace("getire ", "getirmek ")

sadeicerik = sadeicerik.replace("getirmemek ", "getirmemek ")
sadeicerik = sadeicerik.replace("getirmed ", "getirmemek ")
sadeicerik = sadeicerik.replace("getirmey ", "getirmemek ")
sadeicerik = sadeicerik.replace("getirmemes ", "getirmemek ")

sadeicerik = sadeicerik.replace("edilmek ", "edilmek ")
sadeicerik = sadeicerik.replace("edild ", "edilmek ")
sadeicerik = sadeicerik.replace("edilec ", "edilmek ")
sadeicerik = sadeicerik.replace("ediler ", "edilmek ")
sadeicerik = sadeicerik.replace("edilen ", "edilmek ")
sadeicerik = sadeicerik.replace("edili ", "edilmek ")
sadeicerik = sadeicerik.replace("edilmi ", "edilmek ")
sadeicerik = sadeicerik.replace("edilmes ", "edilmek ")

sadeicerik = sadeicerik.replace("edilmemek ", "edilmemek ")
sadeicerik = sadeicerik.replace("edilmed ", "edilmemek ")
sadeicerik = sadeicerik.replace("edilmiy ", "edilmemek ")
sadeicerik = sadeicerik.replace("edilem ", "edilmemek ")
sadeicerik = sadeicerik.replace("edilmey ", "edilmemek ")

```

```
sadeicerik = sadeicerik.replace("etmek ", "etmek ")
sadeicerik = sadeicerik.replace("ediy ", "etmek ")
sadeicerik = sadeicerik.replace("ederi ", "etmek ")
sadeicerik = sadeicerik.replace("edere ", "etmek ")
sadeicerik = sadeicerik.replace("ederl ", "etmek ")
sadeicerik = sadeicerik.replace("edece ", "etmek ")
sadeicerik = sadeicerik.replace("edebil ", "etmek ")
sadeicerik = sadeicerik.replace("etmel ", "etmek ")
sadeicerik = sadeicerik.replace("edipt ", "etmek ")
sadeicerik = sadeicerik.replace("edipd ", "etmek ")
sadeicerik = sadeicerik.replace("etmeli ", "etmek ")
sadeicerik = sadeicerik.replace("edişi ", "etmek ")
sadeicerik = sadeicerik.replace("edişl ", "etmek ")
sadeicerik = sadeicerik.replace("etti ", "etmek ")
sadeicerik = sadeicerik.replace("etmiş ", "etmek ")
```

```
sadeicerik = sadeicerik.replace("etmemek ", "etmemek ")
sadeicerik = sadeicerik.replace("etmemi ", "etmemek ")
sadeicerik = sadeicerik.replace("etmemel ", "etmemek ")
sadeicerik = sadeicerik.replace("etmemes ", "etmemek ")
sadeicerik = sadeicerik.replace("etmey ", "etmemek ")
sadeicerik = sadeicerik.replace("etmiy ", "etmemek ")
sadeicerik = sadeicerik.replace("etmed ", "etmemek ")
sadeicerik = sadeicerik.replace("edem ", "etmemek ")
```

```
sadeicerik = sadeicerik.replace("bırakmak ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakt ", "bırakmak ")
sadeicerik = sadeicerik.replace("bıraka ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakı ", "bırakmak ")
sadeicerik = sadeicerik.replace("bıraks ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakc ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakç ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakmal ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakmas ", "bırakmak ")
sadeicerik = sadeicerik.replace("bırakmış ", "bırakmak ")
```

```
sadeicerik = sadeicerik.replace("bırakmad ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmamas ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmamal ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmad ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmıy ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmay ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmas ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmıc ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmamış ", "bırakmamak ")
sadeicerik = sadeicerik.replace("bırakmamak ", "bırakmamak ")
```

```
sadeicerik = sadeicerik.replace("vermek ", "vermek ")
sadeicerik = sadeicerik.replace("verd ", "vermek ")
sadeicerik = sadeicerik.replace("vermiş ", "vermek ")
sadeicerik = sadeicerik.replace("verip ", "vermek ")
```

```
sadeicerik = sadeicerik.replace("verer ", "vermek ")
sadeicerik = sadeicerik.replace("veren ", "vermek ")
sadeicerik = sadeicerik.replace("verec ", "vermek ")
sadeicerik = sadeicerik.replace("versi ", "vermek ")
sadeicerik = sadeicerik.replace("verse ", "vermek ")
sadeicerik = sadeicerik.replace("veriy ", "vermek ")
sadeicerik = sadeicerik.replace("verir ", "vermek ")
sadeicerik = sadeicerik.replace("verebil ", "vermek ")
sadeicerik = sadeicerik.replace("verilmiş ", "vermek ")
```

```
sadeicerik = sadeicerik.replace("vermemek ", "vermemek ")
sadeicerik = sadeicerik.replace("vermed ", "vermemek ")
sadeicerik = sadeicerik.replace("vermey ", "vermemek ")
sadeicerik = sadeicerik.replace("vermemi ", "vermemek ")
sadeicerik = sadeicerik.replace("vermes ", "vermemek ")
sadeicerik = sadeicerik.replace("vermemes ", "vermemek ")
sadeicerik = sadeicerik.replace("vermiy ", "vermemek ")
sadeicerik = sadeicerik.replace("veremed ", "vermemek ")
sadeicerik = sadeicerik.replace("vermeme ", "vermemek ")
sadeicerik = sadeicerik.replace("vermemekk ", "vermemek ")
sadeicerik = sadeicerik.replace("verilmed ", "vermemek ")
sadeicerik = sadeicerik.replace("verilmiy ", "vermemek ")
sadeicerik = sadeicerik.replace("verilmes ", "vermemek ")
sadeicerik = sadeicerik.replace("vermemel ", "vermemek ")
sadeicerik = sadeicerik.replace("verilmem ", "vermemek ")
sadeicerik = sadeicerik.replace("verilmes ", "vermemek ")
```

```
sadeicerik = sadeicerik.replace("geçeb ", "geçmek ")
sadeicerik = sadeicerik.replace("geçti ", "geçmek ")
sadeicerik = sadeicerik.replace("geçmel ", "geçmek ")
sadeicerik = sadeicerik.replace("geçen ", "geçmek ")
sadeicerik = sadeicerik.replace("geçer ", "geçmek ")
sadeicerik = sadeicerik.replace("geçip ", "geçmek ")
sadeicerik = sadeicerik.replace("geçiş ", "geçmek ")
sadeicerik = sadeicerik.replace("geçil ", "geçmek ")
sadeicerik = sadeicerik.replace("geçirmiş ", "geçmek ")
sadeicerik = sadeicerik.replace("geçiri ", "geçmek ")
sadeicerik = sadeicerik.replace("geçire ", "geçmek ")
sadeicerik = sadeicerik.replace("geçirme ", "geçmek ")
sadeicerik = sadeicerik.replace("geçirmek ", "geçmek ")
sadeicerik = sadeicerik.replace("geçmes ", "geçmek ")
sadeicerik = sadeicerik.replace("geçmel ", "geçmek ")
sadeicerik = sadeicerik.replace("geçmiş ", "geçmek ")
sadeicerik = sadeicerik.replace("geçmek ", "geçmek ")
sadeicerik = sadeicerik.replace("geçme ", "geçmek ")
```

```
sadeicerik = sadeicerik.replace("geçeme ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçem ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmiy ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmey ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmed ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmemi ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmiyor ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçirmem ", "geçmemek ")
```

sadeicerik = sadeicerik.replace("geçirmiy ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçirmey ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçirmed ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçirmeme ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmemek ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmemel ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmemes ", "geçmemek ")
sadeicerik = sadeicerik.replace("geçmeme ", "geçmemek ")

sadeicerik = sadeicerik.replace("çıkılmamak ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkılmamas ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkılmıy ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkılmad ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkılmay ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkılmamam ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkılmamı ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartılmam ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartılmıy ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartılmad ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkaram ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartılmam ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartılmıy ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartılmad ", "çıkılmamak ")
sadeicerik = sadeicerik.replace("çıkartam ", "çıkılmamak ")

sadeicerik = sadeicerik.replace("çıkılmak ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkılmış ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkıac ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkab ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkıy ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkıktı ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkımas ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkırıla ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkırlı ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkartılmış ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkartıld ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkartıla ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkartılı ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkartılmış ", "çıkılmak ")
sadeicerik = sadeicerik.replace("çıkartıld ", "çıkılmak ")

sadeicerik = sadeicerik.replace("olmamak ", "olmamak ")
sadeicerik = sadeicerik.replace("olmad ", "olmamak ")
sadeicerik = sadeicerik.replace("olmay ", "olmamak ")
sadeicerik = sadeicerik.replace("olmuy ", "olmamak ")
sadeicerik = sadeicerik.replace("olama ", "olmamak ")
sadeicerik = sadeicerik.replace("olmamas ", "olmamak ")

sadeicerik = sadeicerik.replace("olmak ", "olmak ")
sadeicerik = sadeicerik.replace("oldu ", "olmak ")
sadeicerik = sadeicerik.replace("olac ", "olmak ")
sadeicerik = sadeicerik.replace("olab ", "olmak ")
sadeicerik = sadeicerik.replace("olan ", "olmak ")
sadeicerik = sadeicerik.replace("olmuş ", "olmak ")

```

sadeicerik = sadeicerik.replace("oluy ", "olmak ")
sadeicerik = sadeicerik.replace("olur ", "olmak ")
sadeicerik = sadeicerik.replace("ols ", "olmak ")

sadeicerik = sadeicerik.replace("göndermek ", "göndermek ")
sadeicerik = sadeicerik.replace("gönderiy ", "göndermek ")
sadeicerik = sadeicerik.replace("gönderil ", "göndermek ")
sadeicerik = sadeicerik.replace("gönderiş ", "göndermek ")
sadeicerik = sadeicerik.replace("gönderd ", "göndermek ")
sadeicerik = sadeicerik.replace("göndermel ", "göndermek ")
sadeicerik = sadeicerik.replace("göndermiş ", "göndermek ")
sadeicerik = sadeicerik.replace("göndere ", "göndermek ")

sadeicerik = sadeicerik.replace("göndermemek ", "göndermemek ")
sadeicerik = sadeicerik.replace("göndermiy ", "göndermemek ")
sadeicerik = sadeicerik.replace("göndermed ", "göndermemek ")
sadeicerik = sadeicerik.replace("göndermey ", "göndermemek ")
sadeicerik = sadeicerik.replace("göndermemel ", "göndermemek ")
sadeicerik = sadeicerik.replace("göndereme ", "göndermemek ")

sadeicerik = sadeicerik.replace("gitmek ", "gitmek ")
sadeicerik = sadeicerik.replace("gidebilm ", "gitmek ")
sadeicerik = sadeicerik.replace("gidebilen ", "gitmek ")
sadeicerik = sadeicerik.replace("gidebild ", "gitmek ")
sadeicerik = sadeicerik.replace("gidec ", "gitmek ")
sadeicerik = sadeicerik.replace("gidiy ", "gitmek ")
sadeicerik = sadeicerik.replace("gitt ", "gitmek ")
sadeicerik = sadeicerik.replace("gidiş ", "gitmek ")
sadeicerik = sadeicerik.replace("gitmiş ", "gitmek ")
sadeicerik = sadeicerik.replace("gitmes ", "gitmek ")
sadeicerik = sadeicerik.replace("giden ", "gitmek ")

sadeicerik = sadeicerik.replace("gitmemek ", "gitmemek ")
sadeicerik = sadeicerik.replace("gidebileme ", "gitmemek ")
sadeicerik = sadeicerik.replace("gideme ", "gitmemek ")
sadeicerik = sadeicerik.replace("gitmiy ", "gitmemek ")
sadeicerik = sadeicerik.replace("gitmey ", "gitmemek ")
sadeicerik = sadeicerik.replace("gitmed ", "gitmemek ")
sadeicerik = sadeicerik.replace("gitmemes ", "gitmemek ")
sadeicerik = sadeicerik.replace("gidemiy ", "gitmemek ")

sadeicerik = sadeicerik.replace("gelmek ", "gelmek ")
sadeicerik = sadeicerik.replace("gelebilm ", "gelmek ")
sadeicerik = sadeicerik.replace("gelebilen ", "gelmek ")
sadeicerik = sadeicerik.replace("gelebild ", "gelmek ")
sadeicerik = sadeicerik.replace("gelec ", "gelmek ")
sadeicerik = sadeicerik.replace("geliy ", "gelmek ")
sadeicerik = sadeicerik.replace("geld ", "gelmek ")
sadeicerik = sadeicerik.replace("geliş ", "gelmek ")
sadeicerik = sadeicerik.replace("gelmiş ", "gelmek ")
sadeicerik = sadeicerik.replace("gelmes ", "gelmek ")
sadeicerik = sadeicerik.replace("gelen ", "gelmek ")

```

```
sadeicerik = sadeicerik.replace("gelmemek ", "gelmemek ")
sadeicerik = sadeicerik.replace("gelmemes ", "gelmemek ")
sadeicerik = sadeicerik.replace("gelebileme ", "gelmemek ")
sadeicerik = sadeicerik.replace("geleme ", "gelmemek ")
sadeicerik = sadeicerik.replace("gelmiy ", "gelmemek ")
sadeicerik = sadeicerik.replace("gelmey ", "gelmemek ")
sadeicerik = sadeicerik.replace("gelmed ", "gelmemek ")
sadeicerik = sadeicerik.replace("gelemiy ", "gelmemek ")
```

```
sadeicerik = sadeicerik.replace("aramak ", "aramak ")
sadeicerik = sadeicerik.replace("aramal ", "aramak ")
sadeicerik = sadeicerik.replace("aradı ", "aramak ")
sadeicerik = sadeicerik.replace("arıy ", "aramak ")
sadeicerik = sadeicerik.replace("aray ", "aramak ")
sadeicerik = sadeicerik.replace("aramı ", "aramak ")
```

```
sadeicerik = sadeicerik.replace("açmamak ", "açmamak ")
sadeicerik = sadeicerik.replace("açmıy ", "açmamak ")
sadeicerik = sadeicerik.replace("açmad ", "açmamak ")
sadeicerik = sadeicerik.replace("açmay ", "açmamak ")
sadeicerik = sadeicerik.replace("açılmıy ", "açılmamak ")
sadeicerik = sadeicerik.replace("açılmad ", "açılmamak ")
```

```
for i in range(0, 50):
```

```
    sadeicerik = str(sadeicerik).replace(" ", " ")
```

```
if sadeicerik[0] == " ":
```

```
    sadeicerik = str(sadeicerik).replace(" ", "", 1)
```

```
if sadeicerik[1] == " ":
```

```
    sadeicerik = str(sadeicerik).replace(" ", "", 1)
```

```
sadeicerik = sadeicerik.replace("brand kargo", "brand")
```

```
sadeicerik = sadeicerik.replace("brand kargo", "brand")
```

```
sadeicerik = sadeicerik.replace(" müşteri hizmet ", " müşteri hizmetleri ")
```

```
print(sadeicerik)
```

```
sheet1.write(n-1, 0, sadeicerik)
```

```
print( str(n-1) + ".SATIR YAZDIRILIYOR %"+str(int((((n-1)*100)/satirsayisi))+ " tamamlandı...")
```

```
wb.save('excel.xls')
```

```

import xlwt
from openpyxl import Workbook, load_workbook
from xlwt import Workbook

wb = Workbook()
sheet1 = wb.add_sheet('Sheet 1')

dosya_ismi = "acik1"

toplam = open(""+dosya_ismi+".txt", "r", encoding="utf-8")
icerik = toplam.read()

splitWords = icerik.split()
print(splitWords)
print(len(splitWords))

telefondegisken=0
evdedegisken=0
dagitimdegisken=0
temizdegisken=0
kapiyadegisken=0
iadegisken=0
teslimdegisken=0
diyalogdegisken=0
kayipdegisken=0

oran=1
#####"Gecikti veya Dağıtım
Çıkmadı"#####
if 'dağıtım' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'çıkmadı' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'çıkış şube' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'gelmedi' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'beklet' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'süredir' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'gecik' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'bekli' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'hala çıkış' in icerik:
    dagitimdegisken=dagitimdegisken+1
if 'geç ' in icerik:
    dagitimdegisken=dagitimdegisken+1
if (dagitimdegisken*100)/len(splitWords)>oran:
    print("Gecikti veya Dağıtım Çıkmadı")
    print(dagitimdegisken)

```

```
#####"Evde yok notu düşüldü veya
Kapıya Getirilmedi"#####
if 'adresiniz'in icerik:
    evdedegisken=evdedegisken+1
if 'adreste'in icerik:
    evdedegisken=evdedegisken+1
if 'evde'in icerik:
    evdedegisken=evdedegisken+1
if 'not 'in icerik:
    evdedegisken=evdedegisken+1
if 'evime'in icerik:
    evdedegisken=evdedegisken+1
if 'kapım'in icerik:
    evdedegisken=evdedegisken+1
if 'çalmıyor'in icerik:
    evdedegisken=evdedegisken+1
if 'çalmadı'in icerik:
    evdedegisken=evdedegisken+1
if 'çalmadan'in icerik:
    evdedegisken=evdedegisken+1
if 'geldik'in icerik:
    evdedegisken=evdedegisken+1
if 'yoktunuz'in icerik:
    evdedegisken=evdedegisken+1
if 'kağı'in icerik:
    evdedegisken=evdedegisken+1
if 'daire'in icerik:
    evdedegisken=evdedegisken+1
if 'bina'in icerik:
    evdedegisken=evdedegisken+1
if 'aşağı'in icerik:
    evdedegisken=evdedegisken+1
if 'inmem'in icerik:
    evdedegisken=evdedegisken+1
if 'yukarı'in icerik:
    evdedegisken=evdedegisken+1
if 'kapıya'in icerik:
    evdedegisken=evdedegisken+1
if(evdedegisken*100)/len(splitWords)>oran:
    print("Evde yok notu düşüldü veya Kapıya Getirilmedi")
    print(evdedegisken)
#####"Telefonlara Cevap
Verilmedi"#####
if 'telefonu'in icerik:
    telefondegisken=telefondegisken+1
if 'telefona'in icerik:
    telefondegisken=telefondegisken+1
if 'telefonlar'in icerik:
    telefondegisken=telefondegisken+1
if 'açmadı'in icerik:
    telefondegisken=telefondegisken+1
if 'cevap'in icerik:
    telefondegisken=telefondegisken+1
if 'açmıyor'in icerik:
```

```

    telefondegisken=telefondegisken+1
if 'ulařılamıyor'ın icerik:
    telefondegisken=telefondegisken+1
if 'ulařamı'ın icerik:
    telefondegisken=telefondegisken+1
if 'aan yok'ın icerik:
    telefondegisken=telefondegisken+1
if(telefondegisken*100)/len(splitWords)>oran:
    print("Telefonlara Cevap Verilmedi")
    print(telefondegisken)
#####İade
Süreci"#####
if 'geri gön'ın icerik:
    iadedegisken=iadedegisken+1
if 'iade'ın icerik:
    iadedegisken=iadedegisken+1
if 'geri yolla'ın icerik:
    iadedegisken=iadedegisken+1
if(iadedegisken*100)/len(splitWords)>oran:
    print("İade Süreci")
    print(iadedegisken)
#####Teslim Alınmadı veya Teslim
Edilmedi"#####
if 'teslimat yapılm'ın icerik:
    teslimdegisken=teslimdegisken+1
if 'teslimatı yapılm'ın icerik:
    teslimdegisken=teslimdegisken+1
if 'teslimat yapma'ın icerik:
    teslimdegisken=teslimdegisken+1
if 'teslim edilme'ın icerik:
    teslimdegisken=teslimdegisken+1
if 'teslim alınma'ın icerik:
    teslimdegisken=teslimdegisken+1
if 'teslim etme'ın icerik:
    teslimdegisken=teslimdegisken+1
if 'teslim alma'ın icerik:
    teslimdegisken=teslimdegisken+1
if(teslimdegisken*100)/len(splitWords)>oran:
    print("Teslim Alınmadı veya Teslim Edilmedi")
    print(teslimdegisken)

```

```
#####"Kötü Diyalog Veya Saygısız
Tutum"#####
if 'saygı'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'huysuz'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'laubali'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'lakayı'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'bağır'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'keyfiye'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'terbiye'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if 'azar'ın icerik:
    diyalogdegisken=diyalogdegisken+1
if(diyalogdegisken*100)/len(splitWords)>oran:
    print("Kötü Diyalog Veya Saygısız Tutum")
    print(diyalogdegisken)
#####"Hasarlı veya Kayıp
Paket"#####
if 'kayıp'ın icerik:
    kayipdegisken=kayipdegisken+1
if 'kayb'ın icerik:
    kayipdegisken=kayipdegisken+1
if 'hasar'ın icerik:
    kayipdegisken=kayipdegisken+1
if 'parça'ın icerik:
    kayipdegisken=kayipdegisken+1
if(kayipdegisken*100)/len(splitWords)>oran:
    print("Hasarlı veya Kayıp Paket")
    print(kayipdegisken)
#####"Hijyen Kurallarına
Uyulmadı"#####
if 'hijyen'ın icerik:
    temizdegisken=temizdegisken+1
if 'maske'ın icerik:
    temizdegisken=temizdegisken+1
if 'temiz'ın icerik:
    temizdegisken=temizdegisken+1
if 'titiz'ın icerik:
    temizdegisken=temizdegisken+1
if(temizdegisken*100)/len(splitWords)>oran:
    print("Hijyen Kurallarına Uyulmadı")
    print(temizdegisken)
#####
#####
```

```

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

from sklearn.feature_extraction.text import HashingVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.multiclass import OneVsRestClassifier
from sklearn.pipeline import Pipeline

train = pd.read_excel('input/train.xls')
sikayetsayisi=len(train)
yontem=TfidfVectorizer

print(train.info())
print(train.describe())
print(train.icerik.head())

#creating x and y
x=train.loc[:, 'icerik']

z=[0,1,2,3,4,5]
z[0]=train.loc[:, 'Gecikti veya Dağıtım Çıkmadı']
z[1]=train.loc[:, 'Evde yok notu düşüldü veya Kapıya Getirilmedi']
z[2]=train.loc[:, 'Telefonlara Cevap Verilmedi']
z[3]=train.loc[:, 'İade Süreci']
z[4]=train.loc[:, 'Teslim Alınmadı veya Teslim Edilmedi']
z[5]=train.loc[:, 'Hijyen Kurallarına Uyulmadı']

y = train.drop(['id', 'icerik'], axis=1)

tks = '[A-Za-z0-9]+(?:\\s+)'

pl = Pipeline([
    ('vec', yontem(token_pattern = tks)),
    ('clf', OneVsRestClassifier(LogisticRegression()))
])

# Fit to the training data
pl.fit(x,y)

test = pd.read_excel('input/train.xls')
test.info()
#1 missing value

test = test.fillna("")
#predicting
predictions = pl.predict_proba(test.icerik)

```

```

toplam=[0,0,0,0,0,0]
ortalama=[0,0,0,0,0,0]
for sutun in range (0,len(z)):
    for i in range (0,len(predictions)):
        toplam[sutun]=toplam[sutun]+predictions[i][sutun]
    ortalama[sutun]=toplam[sutun]/len(predictions)

"""
for i in range (0 , len(predictions)):
    for j in range (0,len(z)):
        if predictions[i][j] < ortalama[j]:
            predictions[i][j] = 0
        else:
            predictions[i][j] = 1
"""

# Format predictions in DataFrame: prediction_df
prediction_df = pd.DataFrame(columns=y.columns, index=test.id, data=predictions)

prediction_df.to_excel('predictions.xls')

degisken=0
toplamtahmin=0

for i in range (0 , sikayetsayisi):
    for j in range (0,len(z)):
        toplamtahmin=toplamtahmin+1
        if predictions[i][j] == z[j][i]:
            degisken=degisken+1

yontem=str(yontem)
if "CountVectorizer" in yontem:
    yontem = str(yontem).replace(yontem, "CountVectorizer")
if "TfidfVectorizer" in yontem:
    yontem = str(yontem).replace(yontem, "TfidfVectorizer")
if "HashingVectorizer" in yontem:
    yontem = str(yontem).replace(yontem, "HashingVectorizer")

if len(test)==len(train):
    print("\n"+yontem+" yöntemi ile train edilen "+str(sikayetsayisi)+" test verisinde yapılan toplam " +
          str(toplamtahmin) + " tahminden " + str(degisken) + " kadarı doğrudur")
    print("doğruluk oranı= %"+ str((100*degisken)/toplamtahmin))
else:
    print("\n"+str(len(test))+ " adet veride tahminleme işlemi "+yontem+" yöntemi ile gerçekleştirilmiştir.")

```

ÖZGEÇMİŞ

Adı Soyadı : Tolga KUYUCUK

ÖĞRENİM DURUMU

Derece	Eğitim Birimi	Mezuniyet Yılı
Yüksek Lisans	Sakarya Üniversitesi / Fen Bilimleri Enstitüsü / Bilişim Sistemleri Mühendisliği	2022
Lisans	İstanbul Üniversitesi / Açık ve Uzaktan Eğitim Fakültesi / Yönetim Bilişim Sistemleri	Devam ediyor
Lisans	Ankara Üniversitesi / Dil, Tarih ve Coğrafya Fakültesi / Antropoloji	Devam ediyor
Lisans	Sakarya Üniversitesi / Bilgisayar ve Bilişim Bilimleri Fakültesi / Bilgisayar Mühendisliği	2017
Ön Lisans	Anadolu Üniversitesi / Açıköğretim Fakültesi Web Tasarımı ve Kodlama	2021

İŞ DENEYİMİ

Yıl	Yer	Görev
2021-2022	7Bits Омск /Rusya	ML Mühendisi
2019-2020	Türk Hava Kuvvetleri	Mühendis Subay
2018-2018	Ubersem Berlin/Almanya	Python Geliştirici
2017-2018	UIZ Berlin/Almanya	Android Geliştirici
2017-2017	Yapı Merkezi Holding	Yazılım Stajyeri
2016-2016	Turkticaret.Net A.Ş.	Donanım Stajyeri

YABANCI DİL

İngilizce, Almanca, Rusça

ESERLER (makale, bildiri, proje vb.)

1. Pandemi Döneminde Kargo Firmalarına Yapılan Şikayetlerin Makine Öğrenmesiyle Sınıflandırılması **(bildiri)**
2. Mobil Telekomünikasyon Operatörlerine Şikayetlerin Metin Madenciliği Yöntemi ile Analizi **(bildiri)**

